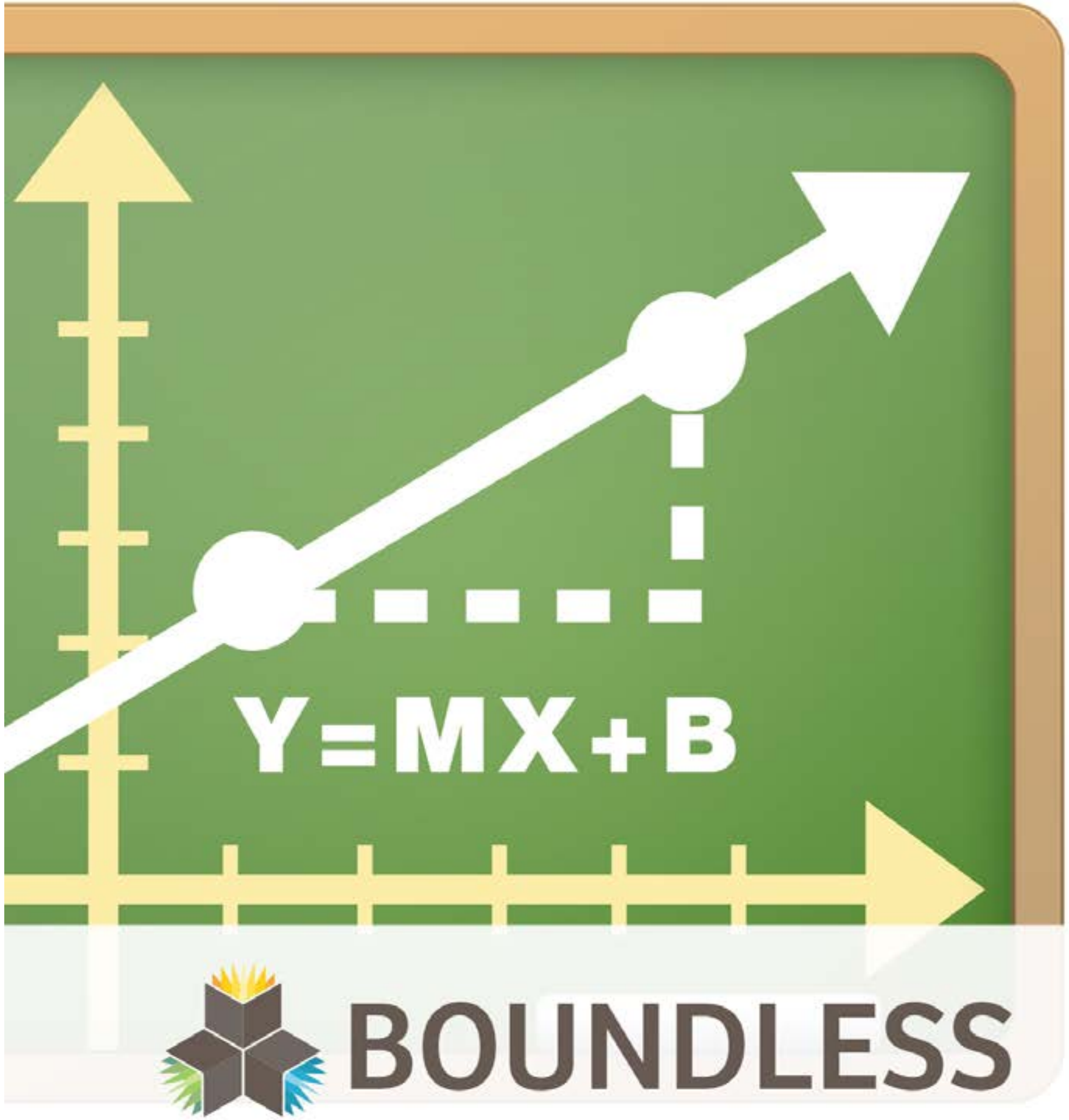
Boundless

Algebra



Algebra

BOUNDLESS

Boundless is better than your assigned textbook.

We create our textbooks by finding the best content from open educational libraries, government resources, and other free learning sites. We then tie it all together with our proprietary process, resulting in great textbooks.

Stop lugging around heavy, expensive, archaic textbooks. Get your Boundless alternative today and see why students at thousands of colleges and universities are getting smart and going Boundless.

Boundless goes beyond a traditional textbook.

Way beyond.

Instant search

Chapters, key term definitions, and anything else at your fingertips.

SmartNotes

It's like your professor summarized the readings for you.

Quizzes

When you feel ready, you can quiz yourself to see how much you know.

Flashcards

Flashcards are a great way to study key terms, concepts and more.

Highlights

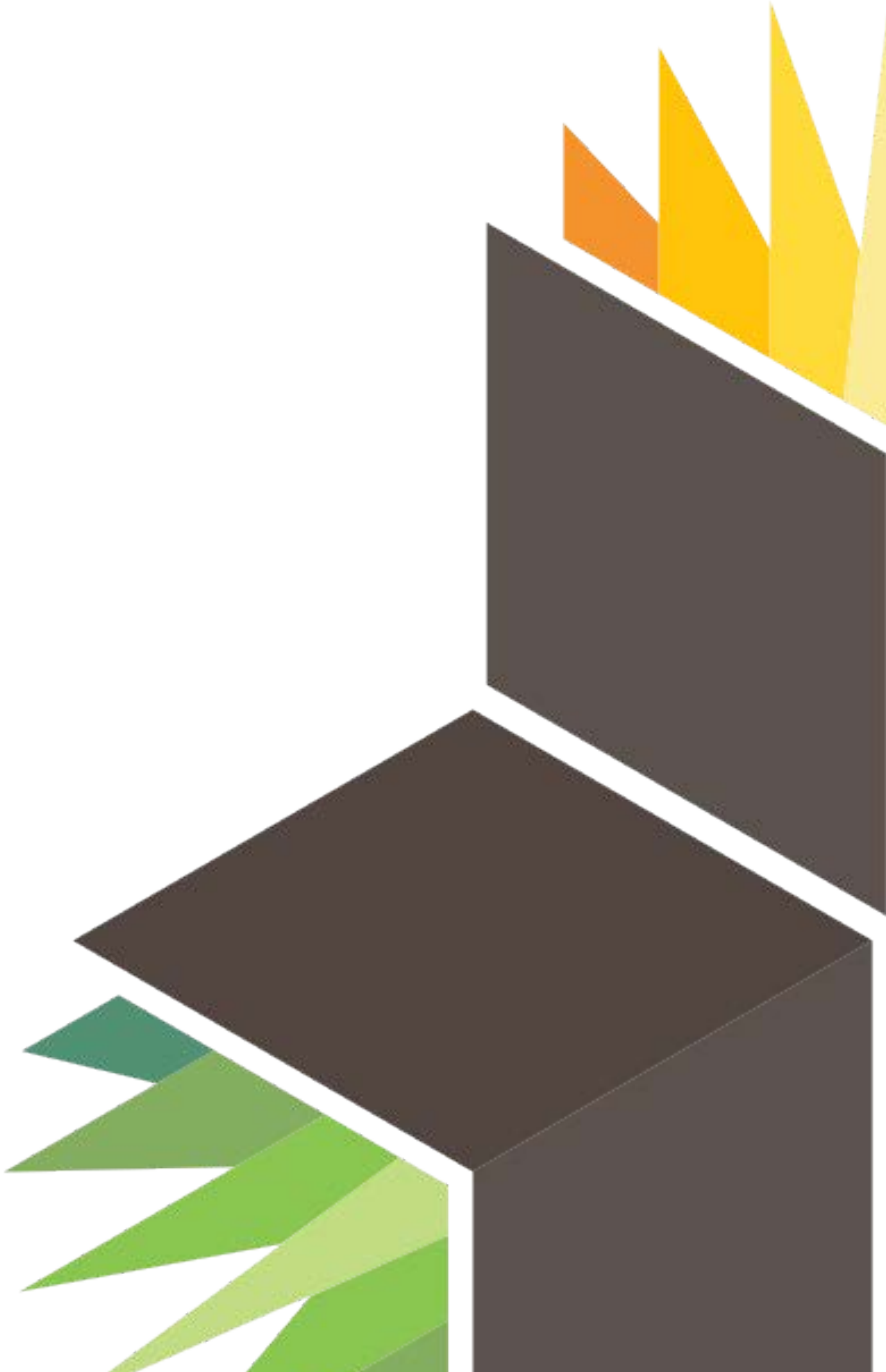
Highlight key points and key terms so you can come back to them later.

Notes

Add notes to your highlights to make them even more meaningful.

Chapter 1

The Building Blocks of Algebra



https://www.boundless.com/algebra/the-building-blocks-of-algebra/

Real Numbers

Real Numbers: Basic Operations

Interval Notation

Equations, Inequalities, Properties

Introduction to Absolute Value

https://www.boundless.com/algebra/the-building-blocks-of-algebra/real-numbers/

Real Numbers: Basic Operations

The basic arithmetic operations for real numbers are addition, subtraction, multiplication, and division.

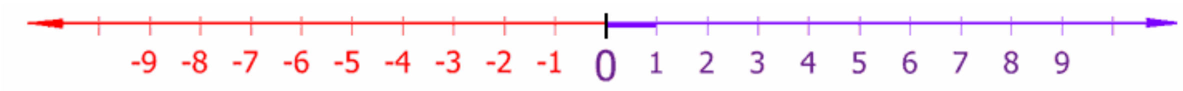
KEY POINTS

- A real number is a value that represents a quantity along a continuous line. Real numbers can be thought of as points on an infinitely long line called the number line or real line.
- The basic arithmetic operations for real numbers are addition, subtraction, multiplication, and division.
- Arithmetic operations are performed according to a specific hierarchy or order, not from left to right.

A real number is a value that represents a quantity along a continuous line. The real numbers include all the **rational numbers**, such as the integer -5 and the fraction $4/3$, and all the irrational numbers, such as $\sqrt{2}$. Real numbers can be thought of as points on an infinitely long line called the number line (real line), where the points corresponding to integers are equally spaced as shown in *Figure 1.1*.

The basic arithmetic operations for real numbers are addition, subtraction, multiplication, and division. Arithmetic operations are performed according to a specific hierarchy or order, not from left to right.

Figure 1.1 Real Numbers



Real numbers can be thought of as points on an infinitely long number line.

Addition and Subtraction

Addition is the basic operation of arithmetic. In its simplest form, addition combines two numbers into a single number. Adding more than two numbers can be viewed as repeated addition; this procedure is known as summation and includes ways to add infinitely many numbers in an infinite series. Addition is **commutative** and **associative**, so the order in which the terms are added does not affect their sum. The identity element of addition is 0; that is, adding zero to any number yields that same number.

Subtraction is the inverse of addition; it finds the difference between two numbers. As such, taking a number x, adding b to it and subsequently subtracting b from it affords the same number x. Subtraction is neither commutative nor associative.

Multiplication and Division

Multiplication also combines two numbers into a single number, the product. Multiplication is best viewed as a simplification of many additions. For example the product of x and y is the sum of x written out y times.

Multiplication is commutative and associative, and its identity is 1. That is, multiplying any number by 1 yields that same number. Division is the inverse of multiplication. Thus, taking a number x and multiplying it by b and then dividing it by b results in the same number x. Division is neither commutative nor associative.

Interval Notation

Intervals notation uses parentheses and brackets to describe sets of real numbers and their endpoints.

KEY POINTS

- A real interval is a set of real numbers with the property that any number that lies between two numbers in the set is also included in the set.
- The interval of numbers between a and b, including a and b, is often denoted [a, b]. The two numbers are called the endpoints of the interval.
- To indicate that one of the endpoints is to be excluded from the set, the corresponding square bracket can be either replaced with a parenthesis, or reversed. One may use an infinite endpoint to indicate that there is no bound in that direction.
- An open interval does not include its endpoints, and is indicated with parentheses. A closed interval includes its endpoints, and is denoted with square brackets.

A real **interval** is a set of real numbers with the property that any number that lies between two numbers in the set is also included in the set. For example, the set of all numbers x satisfying $0 \le x \le 1$ is an interval which contains 0 and 1, as well as all numbers between

them. Other examples of intervals are the set of all real numbers and the set of all negative real numbers.

The interval of numbers between a and b, including a and b, is often denoted [a, b]. The two numbers are called the **endpoints** of the interval. To indicate that one of the endpoints is to be excluded from the set, the corresponding square bracket can be either replaced with a parenthesis, or reversed.

One may use an infinite endpoint to indicate that there is no bound in that direction. For example, $(0, +\infty)$ is the set of all positive real numbers, and $(-\infty, +\infty)$ is the set of real numbers.

An **open interval** does not include its endpoints, and is indicated with parentheses. For example (0,1) means greater than 0 and less than 1. A closed interval includes its endpoints, and is denoted with square brackets. For example [0,1] means greater than or equal to 0 and less than or equal to 1 (*Figure 1.2*).

An interval is said to be left-bounded or right-bounded if there is some real number that is, respectively, smaller than or larger than all its elements. An interval is said to be bounded if it is both left- and right-bounded, and is said to be unbounded otherwise. Intervals that are bounded at only one end are said to be half-bounded. The empty set is bounded, and the set of all reals is the only interval that is unbounded at both ends. Bounded intervals are also commonly known as finite intervals.

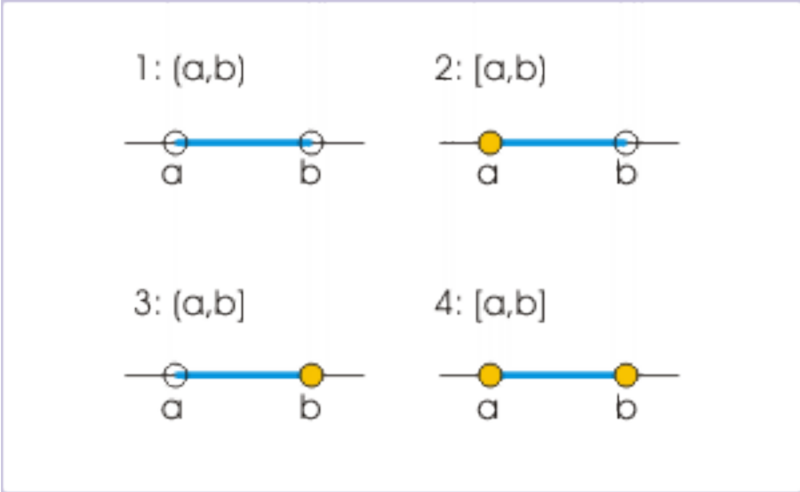


Figure 1.2 Intervals

Representation on real number line.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/real-numbers/interval-notation/

Equations, Inequalities, Properties

An equation states that two expressions are equal, while an inequality relates two different values.

KEY POINTS

- An equation is a mathematical statement that asserts the equality of two expressions.
- An inequality is a relation that holds between two values when they are different.
- The notation $a \neq b$ means that a is not equal to b. It does not say that one is greater than the other, or even that they can be compared in size. If one were to compare the size of the values, the notation $a < b$ means that a is less than b, while the notation $a > b$ means that a is greater than b.

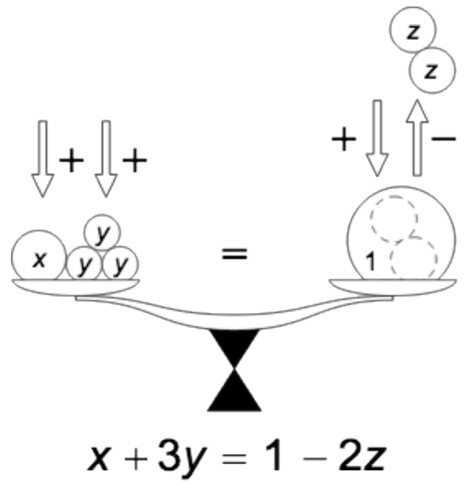


Figure 1.3
Equation as a Balance

Illustration of a simple equation as a balance. x, y, and z are real numbers, analogous to weights.

Equations

An **equation** is a mathematical statement that asserts the equality of two expressions. This is written by placing the expressions on either side of an equals sign (=), for example:

$x + 3 = 5$

asserts that x + 3 is equal to 5 (*Figure 1.3*).

Equations often express relationships between given quantities—the knowns—and quantities yet to be determined—the **unknowns**. By convention, unknowns are denoted by letters at the end of the alphabet, x,y, z, w, ..., while knowns are denoted by letters at the beginning, a, b, c, d, ... The process of expressing the unknowns in terms of the knowns is called solving the equation. In an equation with a single unknown, a value of that unknown for which the equation is true is called a solution or root of the equation. In a set of simultaneous equations, or system of equations, multiple equations are given with multiple unknowns. A solution to the system is an assignment of values to all the unknowns so that all of the equations are true.

Inequalities

An **inequality** is a relation that holds between two values when they are different. The notation $a \neq b$ means that a is not equal to b.

It does not say that one is greater than the other, or even that they can be compared in size.

In either case, a is not equal to b. These relations are known as strict inequalities. To compare the size of the values, there are two types of relations:

- The notation $a < b$ means that a is less than b.
- The notation $a > b$ means that a is greater than b.

The notation $a < b$ may also be read as "a is strictly less than b".

In contrast to strict inequalities, there are two types of inequality relations that are not strict:

- The notation $a \leq b$ means that a is less than or equal to b (or, equivalently, not greater than b, or at most b).
- The notation $a \geq b$ means that a is greater than or equal to b (or, equivalently, not less than b, or at least b)

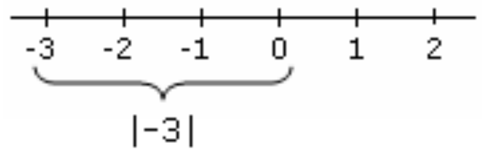
Introduction to Absolute Value

The absolute value may be thought of as the distance of a real number from zero (or the non-negative value without regard to its sign).

KEY POINTS

- The absolute value of a number may be thought of as its distance from zero along the real number line, and more generally, the absolute value of the difference of two real numbers is the distance between them.
- The absolute value $|a|$ of a real number a is the non-negative value of a without regard to its sign. Namely, $|a| = a$ for a positive a, $|a| = -a$ for a negative a, and $|0| = 0$.
- The absolute value of a is always either positive or zero but never negative.

Figure 1.4 Absolute Value

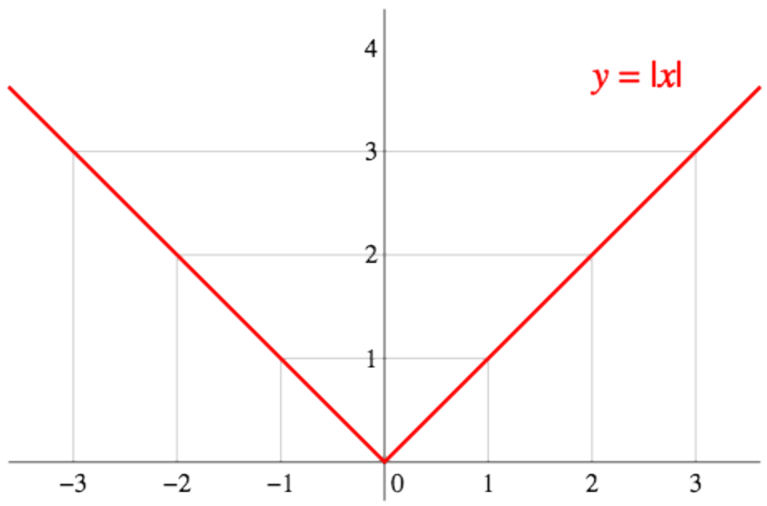


The absolute value of a real number may be thought of as its distance from zero.

The **absolute value** of a number may be thought of as its distance from zero (*Figure 1.4*). In mathematics, the absolute value (or **modulus**) $|a|$ of a real number a is the non-negative value of a

without regard to its sign (*Figure 1.5*). Namely, | a | = a for a positive a, | a | = −a for a negative a, and | 0 | = 0. For example, the absolute value of 3 is 3, and the absolute value of −3 is also 3. The absolute value is closely related to the notions of magnitude, distance, and norm in various mathematical and physical contexts.

Figure 1.5 Absolute Value



The graph of y = |x|. The graph is symmetric for both negative and positive values of x.

Terminology and Notation

The term "absolute value" has been used in this sense since at least 1806 in French and 1857 in English. The notation | a | was introduced by Karl Weierstrass in 1841. Other names for absolute value include "the numerical value" and "the magnitude."

Definition and Properties

For any real number a, the absolute value or modulus of a is denoted by | a | (a vertical bar on each side of the quantity) and is defined as | a | = a for a greater than or equal to 0, and | a | = −a for a < 0. For instance, if a = -3, | -3 | = 3 = −(-3). The double negative yields a positive number. As can be seen from the above definition, the absolute value of a is always either positive or zero but never negative. From an analytic geometry point of view, the absolute value of a real number is that number's distance from zero along the real number line, and more generally, the absolute value of the difference of two real numbers is the distance between them.

Exponents, Scientific Notation, Order of Operations

https://www.boundless.com/algebra/the-building-blocks-of-algebra/exponents-scientific-notation-order-of-operations/

Integer Exponents

An exponent, written b^n, indicated multiplying b times itself n times, so b^3 is b·b·b.

KEY POINTS

- Exponentiation is a mathematical operation, written as b^n, involving two numbers, the base b and the exponent (or index or power) n. When n is a positive integer, exponentiation corresponds to repeated multiplication.
- The expression b^2 = b·b is called the square of b because the area of a square with side-length b is b^2. The expression b^3 = b·b·b is called the cube, because the volume of a cube with side-length b is b^3.
- Some observations may be made about exponents. Any number raised by the exponent 1 is the number itself. Any nonzero number raised by the exponent 0 is 1. These equations do not decide the value of 0^0. Raising 0 by a negative exponent would imply division by 0, so it is undefined.

Exponentiation is a mathematical **operation**, written as b^n, involving two numbers, the **base** b and the **exponent** (or index or power) n. When n is a positive integer, exponentiation corresponds to repeated multiplication; in other words, a product of n factors, each of which is equal to b (the product itself can also be called power) (*Figure 1.7*).

Figure 1.7 Exponent

$$b^n = \underbrace{b \times \cdots \times b}_{n}$$

Taking b to the n power, as shown, is equivalent to multiplying b times itself an n number of times.

Similarly, multiplication by a positive integer corresponds to repeated addition (*Figure 1.6*).

Figure 1.6 Multiplication

$$b \times n = \underbrace{b + \cdots + b}_{n}$$

Exponentiation is related to multiplication in that multiplication of b times n is equivalent to adding b together n number of times.

The exponent is usually shown as a superscript to the right of the base. The exponentiation b^n can be read as: b raised to the n-th power, b raised to the power of n, b raised by the exponent of n, or most briefly as b to the n. Some exponents have their own pronunciation. For example, b^2 is usually read as b squared and b^3 as b cubed. It is also often common to see b^n represented as b^n.

Exponentiation is used pervasively in many fields, including economics, biology, chemistry, physics, and computer science, with applications such as compound interest, population growth, chemical reaction kinetics, wave behavior, and public key cryptography.

Background and Terminology

The expression $b^2 = b \cdot b$ is called the square of b because the area of a square with side-length b is b^2.

The expression $b^3 = b \cdot b \cdot b$ is called the cube, because the volume of a cube with side-length b is b^3.

So 3^2 is pronounced "three squared", and 2^3 is "two cubed."

The exponent says how many copies of the base are multiplied together. For example, $3^5 = 3·3·3·3·3 = 243$. The base 3 appears 5 times in the repeated multiplication, because the exponent is 5. Here, 3 is the base, 5 is the exponent, and 243 is the power or, more specifically, the fifth power of 3, 3 raised to the fifth power, or 3 to the power of 5. The word "raised" is usually omitted, and very often "power" as well, so 3^5 is typically pronounced "three to the fifth" or "three to the five."

Integer Exponents

The exponentiation operation with integer exponents requires only elementary algebra.

Positive Integer Exponents

Formally, powers with positive integer exponents may be defined by the initial condition

$$b^1 = b$$

and the recurrence relation

$$b^{n+1} = b^n \cdot b$$

From the associativity of multiplication, it follows that for any positive integers m and n

$$b^{m+n} = b^m \cdot b^n$$

Arbitrary Integer Exponents

For non-zero b and positive n, the recurrence relation from the previous subsection can be rewritten as

$$b^n = \frac{b^{n+1}}{b}$$

By defining this relation as valid for all integer n and nonzero b, it follows that

$$b^0 = \frac{b^1}{b} = 1$$

$$b^{-1} = \frac{b^0}{b} = \frac{1}{b}$$

and more generally,

$$b^{-n} = \frac{1}{b^n}$$

for any nonzero b and any nonnegative integer n (and indeed any integer n).

The following observations may be made:

- Any number raised by the exponent 1 is the number itself.
- Any nonzero number raised by the exponent 0 is 1; one interpretation of these powers is as empty products.
- These equations do not decide the value of 0^0.
- Raising 0 by a negative exponent would imply division by 0, so it is left undefined.

The identity

$$b^{m+n} = b^m \cdot b^n$$

initially defined only for positive integers m and n, holds for arbitrary integers m and n, with the constraint that m and n must both be positive when b is zero.

Scientific Notation

Scientific notation expresses a number as a$\cdot 10^b$, where a has one digit to the left of the decimal.

KEY POINTS

- Scientific notation is a way of writing numbers that are too big or too small to be conveniently written in decimal form.
- In normalized scientific notation, the exponent b is chosen so that the absolute value of a remains at least one but less than ten ($1 \leq |a| < 10$). Following these rules, 350 would always be written as 3.5×10^2.
- Most calculators present very large and very small results in scientific notation. Because superscripted exponents like 10^7 cannot always be conveniently displayed, the letter E or e is often used to represent "times ten raised to the power of" (which would be written as "x 10^b ").

Standard Form to Scientific Form

Very large numbers such as 43,000,000,000,000,000,000 (the number of different possible configurations of Rubik's cube) and very small numbers such as 0.000000000000000000000340 (the mass of the amino acid tryptophan) are extremely inconvenient to write and read. Such numbers can be expressed more conveniently by writing them as part of a power of 10.

To see how this is done, let us start with a somewhat smaller number such as 2480.

The last form in *Figure 1.8* is called the scientific form of the number. There is one nonzero digit to the left of the decimal point and the absolute value of the exponent on 10 records the number of places the original decimal point was moved to the left. If instead we have a very small number, such as 0.00059, we instead move the decimal place to the right, as in the following:

$$0.00059 = \frac{5.9}{10000} = \frac{5.9}{10^4} = 5.9 \cdot 10^{-4}$$

Figure 1.8 Standard Form Versus Scientific Form

$$\begin{aligned}\underbrace{2480}_{\text{Standard form}} &= 248.0 \times 10^1 \\ &= 24.80 \times 10^2 \\ &= \underbrace{2.480 \times 10^3}_{\text{Scientific form}}\end{aligned}$$

In standard form, the number is written out as you are accustomed to, the ones digit to the farthest to the right (unless there is a decimal), then the tens digit to the left of the ones, and so on. In scientific notation, a number in standard notation with one nonzero digit to the left of the decimal is multiplied by ten to some power, as shown.

There is one nonzero digit to the left of the decimal point and the absolute value of the exponent of 10 records the number of places the original decimal point was moved to the right.

Writing a Number in Scientific Notation

To write a number in **scientific notation**:

- Move the decimal point so that there is one nonzero digit to its left.
- Multiply the result by a power of 10 using an exponent whose absolute value is the number of places the decimal point was moved. Make the exponent positive if the decimal point was moved to the left and negative if the decimal point was moved to the right.

A number written in scientific notation can be converted to standard form by reversing the process described above.

Normalized Scientific Notation

Any given number can be written in the form of $a \times 10^b$ in many ways; for example, 350 can be written as 3.5×10^2 or 35×10^1 or 350×10^0. In normalized scientific notation, the exponent b is chosen so that the absolute value of a remains at least one but less than ten ($1 \le |a| < 10$). Following these rules, 350 would always be written as 3.5×10^2. This form allows easy comparison of two numbers of the same sign in a, as the exponent b gives the number's order of magnitude. In normalized notation, the exponent b is negative for a number with absolute value between 0 and 1 (e.g., negative one half is written as -5×10^{-1}). The 10 and exponent are usually omitted when the exponent is 0. Note that 0 cannot be written in normalized scientific notation since it cannot be expressed as $a \times 10^b$ for any non-zero a. Normalized scientific form is the typical form of expression of large numbers for many fields, except during intermediate calculations or when an unnormalised form, such as engineering notation, is desired. Normalized scientific notation is often called exponential notation—although the latter term is more general and also applies when a is not restricted to the range 1 to 10 (as in engineering notation for instance) and to bases other than 10 (as in $3^{15} \times 2^{20}$).

E Notation

Most calculators and many computer programs present very large and very small results in scientific notation. Because superscripted exponents like 10^7 cannot always be conveniently displayed, the letter E or e is often used to represent "times ten raised to the power of" (which would be written as "x 10^b ") and is followed by the value of the exponent. Note that in this usage the character e is not related to the mathematical constant e or the exponential function e^x (a confusion that is less likely with capital E), and though it stands for

exponent, the notation is usually referred to as (scientific) E notation or (scientific) e notation, rather than (scientific) exponential notation. The use of this notation is not encouraged by publications.

Order of Operations

Order of Operations is a way of evaluating expressions with more than one operation and it governs precedence in mathematical operations.

KEY POINTS

- In order to communicate using mathematical expressions there must be a logical and agreed upon order of operations so that each expression can be written unambiguously.
- The order of operations, or precedence, used throughout mathematics, science, technology and many computer programming languages is expressed here:
  1) terms inside parentheses or brackets
  2) exponents and roots
  3) multiplication and division
  4) addition and subtraction.
- Mnemonics are often used to help students remember the rules. In the U.S. the acronym PEMDAS is common. It stands for Parentheses, Exponents, Multiplication, Division, Addition, Subtraction. PEMDAS is often expanded to "Please Excuse My Dear Aunt Sally".

Order of Operations is a way of evaluating expressions with more than one operation. These rules govern precedence in **mathematical operations**, such as *Figure 1.9*.

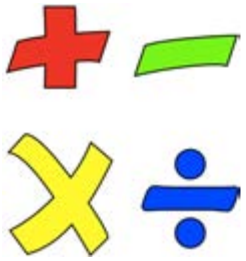


Figure 1.9 Common mathematical operators

The four most common mathematical operators.

For example, when faced with $4 + 2 \cdot 3$, how do you proceed?

There are two apparent options:

$4 + 2 \cdot 3 = (4 + 2) \cdot 3$

$4 + 2 \cdot 3 = 6 \cdot 3$

$4 + 2 \cdot 3 = 18$

OR

$4 + 2 \cdot 3 = 4 + (2 \cdot 3)$

$4 + 2 \cdot 3 = 4 + 6$

$4 + 2 \cdot 3 = 10$

Which one is the correct order of operations in which to solve the problem?

In order to communicate using mathematical expressions there must be a logical and agreed upon order of operations so that each expression can be written unambiguously. For the above example, all mathematicians agree the correct answer is 10. The key question is: what is the order upon which mathematicians have agreed?

The order of operations, or precedence, used throughout mathematics, science, technology, and many computer programming languages is expressed here:

Order of Operations

1. terms inside parentheses or brackets
2. exponents and roots
3. multiplication and division
4. addition and subtraction

Theserules means that if a mathematical expression is preceded by one operator and followed by another, the operator higher on the list should be applied first. The commutative and associative laws of addition and multiplication allow terms to be added in any order and factors to be multiplied in any order, but mixed operations must obey the standard order of operations.

It is helpful to treat division as multiplication by the reciprocal (multiplicative inverse) and subtraction as addition of the opposite (additive inverse). Thus $3/4 = 3 \div 4 = 3 \cdot \frac{1}{4}$; in other words the quotient of 3 and 4 equals the product of 3 and ¼. Also $3 - 4 = 3 + (-4)$; in other words the difference of 3 and 4 equals the sum of positive three and negative four. With this understanding,

think of $1 - 3 + 7$ as the sum of 1, negative 3, and 7, and add in any order: $(1 - 3) + 7 = -2 + 7 = 5$ and in reverse order $(7 - 3) + 1 = 4 + 1 = 5$. The important thing is to keep the negative sign with the 3.

Mnemonics

Mnemonics are often used to help students remember the rules, but the rules taught by the use of acronyms can be misleading. In the United States the acronym PEMDAS is common. It stands for Parentheses, Exponents, Multiplication, Division, Addition, and Subtraction. PEMDAS is often expanded to "Please Excuse My Dear Aunt Sally" with the first letter of each word creating the acronym PEMDAS.

These mnemonics may be misleading when written this way, especially if the user is not aware that multiplication and division are of equal precedence, as are addition and subtraction. Using any of the above rules in the order "addition first, subtraction afterward" would also give the wrong answer.

$10 - 3 + 2$

The correct answer is 9, and not 5, which we get when we add 3 and 2 first to get 5,and then subtract it from 10 to get the final answer of 5, which is best understood by thinking of the problem as the sum of positive ten, negative three, and positive two.

$10 + (-3) + 2$

An alternative way to write the mnemonic is:

P

E

MD

AS

Or, simply as PEMA, where it is taught that multiplication and division inherently share the same precedence and that addition and subtraction inherently share the same precedence. This mnemonic makes the equivalence of multiplication and division and of addition and subtraction clear.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/exponents-scientific-notation-order-of-operations/order-of-operations/

Section 3

Working with Polynomials

https://www.boundless.com/algebra/the-building-blocks-of-algebra/working-with-polynomials/

Polynomials: Introduction, Addition, and Subtration

A polynomial is a finite expression containing constants and variables connected only through basic operations of algebra.

KEY POINTS

- A polynomial is an expression of finite length constructed from variables and constants, using only the operations of addition, subtraction, multiplication, and non-negative integer exponents.
- A polynomial is either zero or can be written as the sum of a finite number of non-zero terms. Each term consists of the product of a constant (called the coefficient of the term) and a finite number of variables (usually represented by letters) raised to whole number powers.
- Polynomials can be added or subtracted using the associative law of addition (grouping all their terms together into a single sum), possibly followed by reordering, and combining of like terms.

A **polynomial** is an expression of finite length constructed from variables and constants, using only the operations of addition, subtraction, multiplication, and non-negative integer exponents.

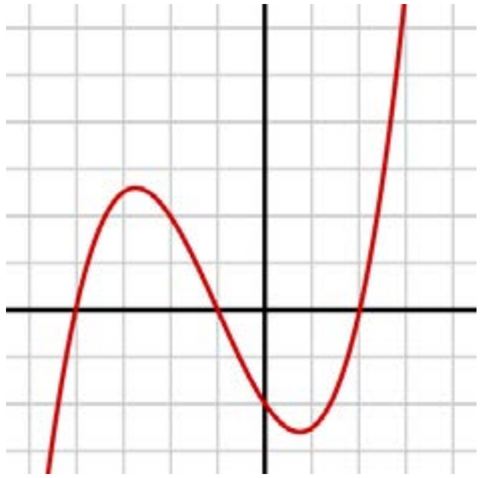


Figure 1.10 **Polynomial**
Graph of a polynomial.

For example, $x^2 - x/4 + 7$ is a polynomial, but $x^2 - 4/x + 7x^{3/2}$ is not. Why? Because its second term involves division by the variable x (4/x), and its third term contains an exponent that is not a non-negative integer (3/2).

Polynomials appear in a wide variety of areas of mathematics and science. For example, they are used to form polynomial equations, which encode a wide range of problems, from elementary word problems to complicated problems in the sciences. They are used to define polynomial functions, which appear in settings ranging from basic chemistry and physics to economics and social science. A graph of a polynomial is shown in *Figure 1.10*.

A polynomial is either zero or can be written as the sum of a finite number of non-zero terms. Each term consists of the product of a constant (called the **coefficient** of the term) and a finite number of variables raised to whole number powers. The value of a variable's exponent is called the **degree** of that variable. The degree of the term is the sum of the degrees of the variables in that term, and the

degree of a polynomial is the largest degree of any one term. If a variable is written without an exponent, you are to assume the degree of that variable is 1.

A constant is a term with no variable or degree.

The commutative law of addition can be used to rearrange terms into any preferred order. In polynomials with one variable, the terms are usually ordered according to degree, either in "descending powers of x", with the term of largest degree first, or in "ascending powers of x".

Terms that are similar, meaning they contain the same variables of the same degree, can be combined by adding the coefficients.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/working-with-polynomials/polynomials-introduction-addition-and-subtration/

Multiplication and Special Products

To multiply two polynomials together, multiply every term of one polynomial by every term of the other polynomial.

KEY POINTS

- To multiply a polynomial by a monomial, multiply every term of the polynomial by the monomial and then add the resulting products together.
- To multiply two polynomials together, multiply every term of one polynomial by every term of the other polynomial.
- Three binomial products occur so frequently in algebra that we designate them as special binomial products. These special products can be shown as the squares of a binomial $(a+b)^2$ and $(a-b)^2$ and as the sum and difference of two terms: (a + b)(a – b).

Multiplication

Multiplying a **polynomial** by a **monomial** is a direct application of the distributive property.

$a(b+c) = ab + ac$

The distributive property suggests the following rule: to multiply a polynomial by a monomial, multiply every term of the polynomial by the monomial and then add the resulting products together. An example is shown in *Figure 1.11*.

Figure 1.11 Polynomial Multiplication

$$3(x+9) = 3 \cdot x + 3 \cdot 9$$
$$= 3x + 27$$

An example of multiplying a polynomial by a monomial.

To multiply a polynomial by a polynomial, we have, by the distributive property:

$(a + b)(c + d) = (a + b)c + (a + b)d = ac + bc + ad + bd$

For convenience, we will use the **commutative** property of addition to write this expression so that the first two terms contain a and the second two contain b:

$(a + b)(c + d) = ac + ad + bc + bd$

This method is commonly called the FOIL method.

- F - First terms
- O - Outer terms
- I - Inner terms
- L - Last terms

Special Products

Three binomial products occur so frequently in algebra that we designate them as special binomial products. These special products can be shown as the squares of a binomial

$(a + b)^2$ and $(a - b)^2$

and as the sum and difference of two terms:

$(a + b)(a - b)$.

There are two simple rules that allow us to easily expand (multiply out) these binomials. They are well worth memorizing, as they will save a lot of time in the future.

To square a binomial:

1. Square the first term.
2. Take the product of the two terms and double it.
3. Square the last term.
4. Add the three results together.

$(a + b)^2 = a^2 + 2ab + b^2$

$(a - b)^2 = a^2 - 2ab + b^2$

To expand the sum and difference of two terms:

1. Square the first term and square the second term.
2. Subtract the square of the second term from the square of the first term.

$(a+b)(a-b)=a^2-b^2$

These are just simplified rules of FOIL. If you do forget these specific rules, you can go ahead and use the FOIL method.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/working-with-polynomials/multiplication-and-special-products/

Dividing Polynomials

Polynomial long division is a method for dividing a polynomial by another polynomial of the same or lower degree.

KEY POINTS

- Polynomial long division is an algorithm for dividing a polynomial by another polynomial of the same or lower degree, a generalized version of the familiar arithmetic technique called long division.
- Polynomial long division can be done easily by hand, because it separates an otherwise complex division problem into smaller ones.
- It is easiest to perform polynomial division when you have a term for each degree power, even if one is just a place holder. For example, if you have a polynomial that looks like: $x^3-12x^2+0x-42$.

Polynomial **long division** is an algorithm for dividing a polynomial by another polynomial of the same or lower degree. This method is a generalized version of the familiar arithmetic technique called long division. It can be done easily by hand, because it separates an otherwise complex division problem into smaller ones.

For example, find the quotient and the remainder of the division of $x^3 - 12x^2 - 42$,, the dividend, by x - 3, the **divisor**. The **dividend** is first rewritten as follows: $x^3 - 12x^2 + 0x - 42$

The quotient and remainder can then be determined as follows:

- Divide the first term of the dividend by the highest term of the divisor (meaning the one with the highest power of x, which in this case is x): $x^3 \div x = x^2$.
- Multiply the divisor by the result just obtained (the first term of the eventual quotient): $x^2 \cdot (x - 3) = x^3 - 3x^2$.
- Subtract the product just obtained from the appropriate terms of the original dividend (being careful that subtracting something having a minus sign is equivalent to adding something having a plus sign): $(x^3 - 12x^2) - (x^3 - 3x^2) = -12x^2 + 3x^2 = -9x^2$.
- Repeat the previous three steps.
- Repeat step 4. This time, there is nothing to "pull down".

The calculated polynomial is the quotient, and the number left over (−123) is the remainder:

$x^3 - 12x^2 - 42 = (x - 3)(x^2 - 9x - 27) - 123$

Another example of polynomial long division is shown in *Figure 1.12*.

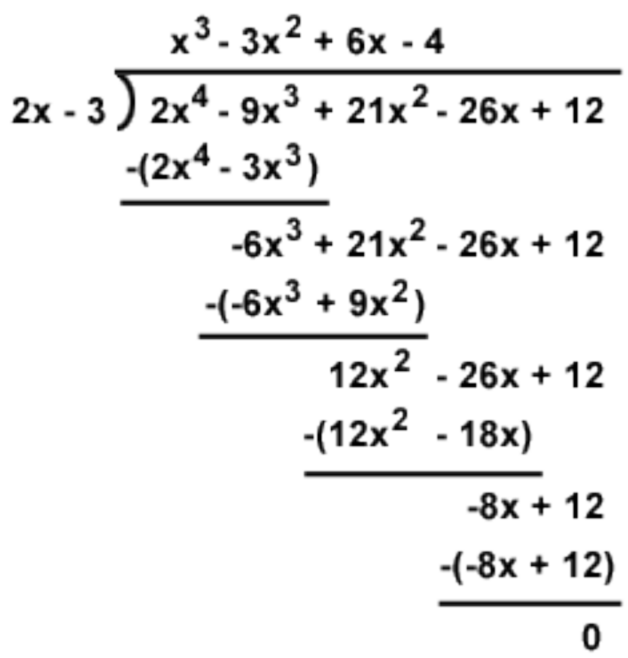


Figure 1.12
Polynomial Long Division

This image shows an example of a polynomial long division.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/working-with-polynomials/dividing-polynomials/

Section 4

Factoring

https://www.boundless.com/algebra/the-building-blocks-of-algebra/factoring/

Greatest Common Factor and Factoring by Grouping

Factoring by grouping divides the terms in a polynomial into groups, which can be factored using the greatest common factor.

KEY POINTS

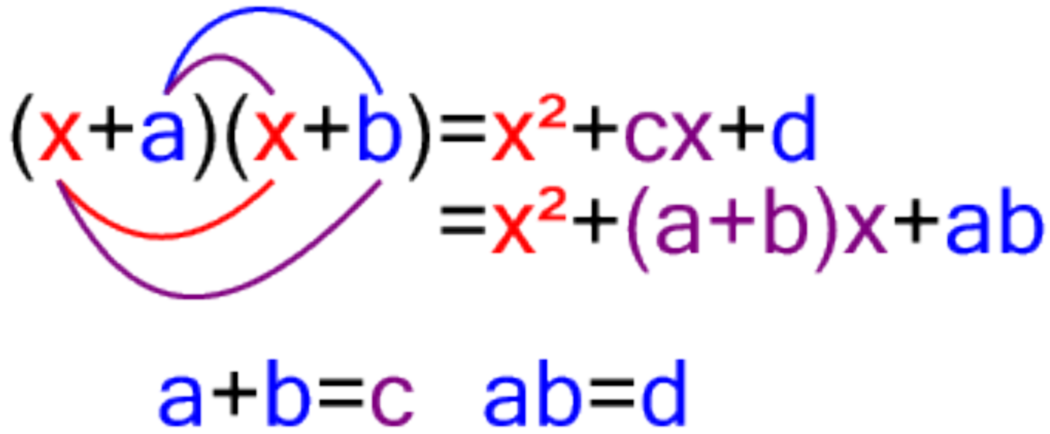
- Factorization or factoring is the decomposition of an object, for example, a number or a polynomial, into a product of other objects, or factors, which when multiplied together give the original.
- The greatest common factor of two polynomials is a polynomial, of the highest possible degree, that evenly divides each of the two original polynomials.
- Factoring by grouping is done by placing the terms in the polynomial into two or more groups, where each group can be factored by a known method. The results of these factorizations can sometimes be combined to make an even more simplified expression.

Factorization or factoring is the decomposition of an object, for example, a number or a **polynomial**, into a product of other objects, or factors, which when multiplied together give the original. As an example, the number 15 factors as 3 × 5, and the polynomial $x^2 - 4$ factors as $(x - 2)(x + 2)$. In all cases, a product of simpler objects is obtained.

The aim of factoring is usually to reduce something to "basic building blocks", such as numbers to prime numbers, or polynomials to irreducible polynomials. Factoring integers is covered by the fundamental theorem of arithmetic and factoring polynomials by the fundamental theorem of algebra.

The opposite of polynomial factorization is expansion, the multiplying together of polynomial factors to an "expanded" polynomial, written as just a sum of terms. The relationship between the two processes can be seen in *Figure 1.13*.

Figure 1.13 Factorization



A visual illustration of the polynomial $x^2 + cx + d = (x + a)(x + b)$ where a plus b equals c and a times b equals d.

Greatest Common Factor

The **greatest common divisor** (GCD), also known as the greatest common factor (GCF), of two or more non-zero integers, is the largest positive integer that divides the numbers without a remainder. For example, the GCD of 8 and 12 is 4. This notion can be extended to polynomials. In algebra, the greatest common divisor of two polynomials is a polynomial, of the highest possible degree, that evenly divides each of the two original polynomials.

Factoring by Grouping

A way to factor some polynomials is factoring by grouping. For those who like algorithms, "factoring by grouping" may be the best way to approach factoring a trinomial, as it takes the guess work out of the process.

Factoring by grouping is done by placing the terms in the polynomial into two or more groups, where each group can be factored by a known method. The results of these factorizations can sometimes be combined to make an even more simplified expression. For example, to factor the polynomial

$4x^2 + 20x + 3yx + 15y$

Group similar terms, $(4x^2 + 20x) + (3yx + 15y)$

Factor out the greatest common factor, $4x(x + 5) + 3y(x + 5)$

Factor out binomial $(x + 5)(4x + 3y)$

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/factoring/greatest-common-factor-and-factoring-by-grouping/

Factoring Trinomials of the Form: $ax^2 + bx + c$; Perfect Squares

The polynomial $ax^2 + bx + c$ can be factored using a variety of methods, including trial and error.

KEY POINTS

- Some trinomials, known as perfect square trinomials, can be factored into two equal binomials.
- We can factor $a^2 - b^2$, the difference of two squares, by finding the terms that produce the perfect squares and substituting these quantities into the factorization form. When using real numbers, there is no factored form for the sum of two squares.
- Perfect square trinomials factor as the square of a binomial. To recognize them, look for whether (1) the first and last terms are perfect squares, and (2) the middle term is divisible by 2, and when halved, equals the product of the terms that when squared produce the first and last terms.

Factoring Trinomials

The polynomial $ax^2 + bx + c$ (like the one graphed below) can be factored using a variety of methods. One such method is trial and error (Figure 1.14).

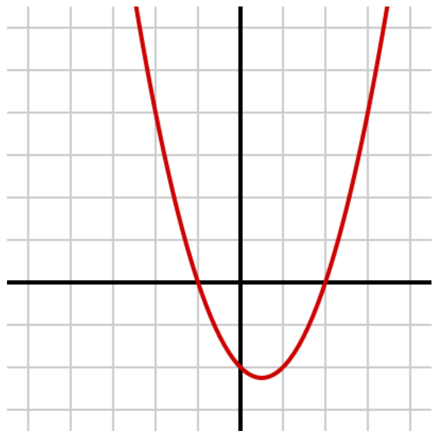


Figure 1.14 Polynomial

Graph of a trinomial $ax^2 + bx + c$.

Ultimately, the **trinomial** should be factored in the form $(px + q)(rx + s)$, where p, q, r, and s are constants, and x is a variable. Using trial and error, we can find values for each of the constants, using the FOIL method to determine whether the constants used produce the trinomial $ax^2 + bx + c$. We know that the product of px and rx must equal a x^2. Additionally, the sum of products px*s and q*rx must equal bx. Finally, the product of q and s must equal c.

Perfect Squares

Some trinomials, known as perfect square trinomials, can be factored into two equal **binomials**. For example:

$a^2 + 2ab + b^2 = (a + b)^2$ and $a^2 - 2ab + b^2 = (a - b)^2$

Perfect square trinomials always factor as the square of a binomial.

To recognize a perfect square trinomial, look for the following features:

1. The first and last terms are perfect squares.
2. The middle term is divisible by 2.

For example, $x^2 - 10x + 25$ can be identified as a perfect square because x^2 is the square of x, and 25 is the square of 5. The middle term (-10x) is divisible by 2 (equalling -5x).

Given that the coefficient of x^2 is 1, we know that the factored form will be $(x + a)(x + b)$, where a and b are to-be-determined coefficients. We need x*b+a*x to equal -10x, and a*b to equal 25. Filling in -5 for a and b, we find a plausible solution that reads (x-5)(x-5), or $(x - 5)^2$. This is a perfect square.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/factoring/factoring-trinomials-of-the-form-ax-2-bx-c-perfect-squares/

Trinomials of the Form: $ax^2 + bx + c$ Where a is not Equal to 1

Two methods for factoring polynomials $ax^2 + bx + c$ are the trial and error method and the collect and discard method.

KEY POINTS

- Two methods for factoring polynomials $ax^2 + bx + c$ are the trial and error method and the collect and discard method.
- For the trial and error method, we look for some factors of the first and last terms. Our goal is to choose the proper combination of factors of the first and last terms that yield the middle term.
- The collect and discard method requires less guessing than the trial and error method.

Factoring Trinomials of the Form $ax^2 + bx + c$

We can easily factor **trinomials** of the form $ax^2 + bx + c$ (*Figure 1.15*) by finding the factors of the constant c that add to the **coefficient** of the **linear** term b, as shown in the following example:

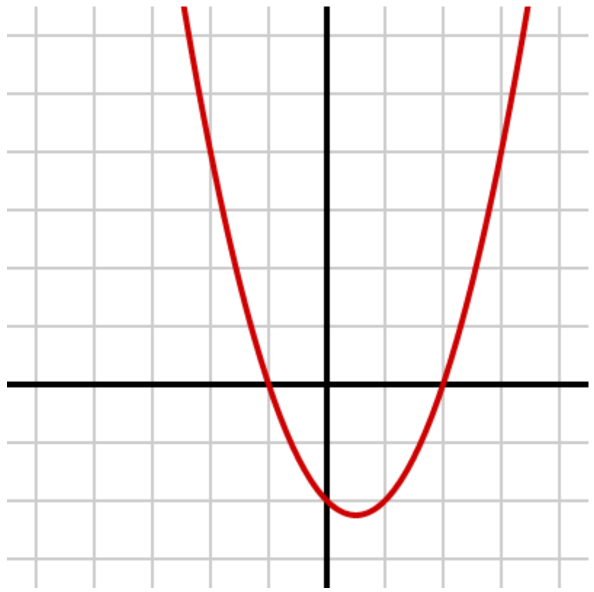


Figure 1.15
Polynomial
A graph of a trinomial $ax^2 + bx + c$ with a = 1.

Factor $ax^2 + 4x + 21$.

The third term of the trinomial is −21. We seek two numbers whose product is −21 and sum is -4. Clearly, the required numbers are −7 and +3.
$4x - 21 = (x - 7)(x + 3)$

The problem of factoring the polynomial , $ax^2 + bx + c, a \neq 1$, becomes more involved. Two methods for factoring these polynomials are the trial and error method and the collect and discard method. Each method produces the same result, and you should select the method you prefer. The trial and error method requires some educated guesses, while the collect and discard method requires less guessing.

Trial and Error Method

For the trial and error method, consider a product. Examining a trinomial, we look for some factors of the first and last terms. Our goal is to choose the proper combination of factors of the first and last terms that yield the middle term. Notice the middle term comes from the sum of the outer and inner products in the multiplication of the two binomials. This fact provides us a way to find the proper combination. Look for the combination that when multiplied and then added yields the middle term.

Collect and Discard Method

For the collect and discard method, consider the polynomial $6x^2 + x - 12$. We begin by identifying a and c. In this case, a = 6 and c = −12. We start out as we would with a=1.
$6x^2 + x - 12 \rightarrow (6x + k)(6x + m)$
where k and m are constants.

Now, compute :
Find the factors of −72 that add to 1, the coefficient of x, the linear term. The factors are 9 and −8. Include these factors in the parentheses.
$6x^2 + x - 12 \rightarrow (6x + 9)(6x - 8)$
But we have included too much. We must eliminate the surplus.

Factor each parentheses.

$6x^2 + x - 12 \rightarrow 3(2x + 3) \cdot 2(3x - 4)$

Discard the factors that multiply to a = 6. In this case, 3 and 2. We are left with the proper factorization.

$6x^2 + x - 12 = (2x + 3)(3x - 4)$

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/factoring/trinomials-of-the-form-ax-2-bx-c-where-a-is-not-equal-to-1/

Special Factorizations and Binomials

By simplifying the FOIL method, you can save a lot of time when factoring for the difference of two square and perfect square trinomials.

KEY POINTS

- We can factor $a^2 - b^2$, the difference of two squares, by finding the terms that produce the perfect squares and substituting these quantities into the factorization form. When using real numbers, there is no factored form for the sum of two squares.
- Perfect square trinomials always factor as the square of a binomial.
- To recognize a perfect square trinomial, look for:
  (1) The first and last terms are perfect squares: $a^2 + 2ab + b^2$
  (2) The middle term is divisible by 2: $a^2 + 2ab + b^2$
  (3)The middle term, once divided by 2, is a product of the square roots of the first and last terms: $a^2 + 2ab + b^2$.

The Difference of Two Squares

Recall that when we multiplied together the two **binomials** (a + b) and (a – b), we obtained the product $a^2 - b^2$: This is obtained by

using the FOIL method, and then adding like terms. Remember, FOIL means First, Outer, Inner, Last.

$$(a+b)(a-b) = (a*a) + (a*-b) + (b*a) + (b*-b)$$

$$= a^2 - ab + ab - b^2 = a^2 - b^2$$

Since we know that $(a+b)(a-b) = a^2 - b^2$, we need only turn the equation around to find the factorization form:

$$a^2 - b^2 = (a+b)(a-b)$$

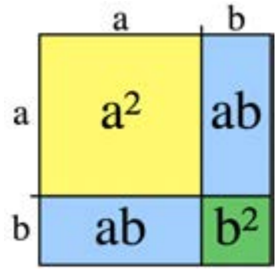
Note: a perfect **square** is a term that is the square of another term. 4 is a perfect square of 2, 9 is a perfect square of 3, a^2 is a perfect square of a

The factorization form says that we can factor $a^2 - b^2$, the difference of two squares, by finding the terms that produce the perfect squares and substituting these quantities into the factorization form. When using real numbers (as we are), there is no factored form for the sum of two squares. That is, using real numbers, $a^2 + b^2$ cannot be factored.

Perfect Square Trinomials

Recall the process of squaring a binomial, which is done using the FOIL method:

Figure 1.16 Perfect Square Trinomial



A visual illustration of $(a+b)^2 = a^2 + 2ab + b^2$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$(a-b)^2 = a^2 - 2ab + b^2$$

Perfect square **trinomials** (*Figure 1.16*) always factor as the square of a binomial. To recognize a perfect square trinomial, look for the following features:

1. The first and last terms are perfect squares: $a^2 + 2ab + b^2$
2. The middle term is divisible by 2: $a^2 + 2ab + b^2$
3. The middle term, once divided by 2, is a product of the square roots of the first and last terms: $a^2 + 2ab + b^2$

In other words, factoring a perfect square trinomial amounts to finding the terms that, when squared, produce the first and last terms of the trinomial, and substituting into one of the formula

$$a^2 + 2ab + b^2 = (a+b)^2$$

$$a^2 - 2ab + b^2 = (a-b)^2$$

EXAMPLES

Factor $x^2 - 16$. This example is pretty straightforward. Let's look at the first term in the expression first: x^2, it is really easy to see that this is a perfect square of x. Now let's look at the second term: 16, this is a perfect square of 4 so now we have x _ 16. The last piece of the puzzle is to find the sign. We can see that in the original expression, that the sign before the 16 is negative. We can try the equation $(x - 4)^2 : (x - 4)(x - 4) = x^2 - 8x + 16$ This is obviously not the answer we are looking for, so let's try another approach. We know that we need the 16 to be negative, and the only way to make that happen is by:(x-4)(x+4). Before submitting anything, always check your answer. When you try this equation, you get: $(x - 4)(x + 4) = x^2 - 4x + 4x - 16 = x^2 - 16$. It works! As you do more examples, you will start to be able to recognize patterns without having to go through so many steps.

Factor $x^2 + 6x + 9$ In the text, we outlined three steps to factoring out a trinomial. Let's start with the first step: 1. The first and last terms are perfect squares: x^2 and 9 are both perfect squares, of x and 3 respectively. Check! 2. The middle term must be divisible by 2:6x/2 = 3x. Check! 3. The middle term, once divided by 2, is a product of the square roots of the first and last terms. Let's take the roots of the first two terms that we found in step 1 and find their product. x times 3 = 3x. Now let's multiply this by 2: 2(3x) = 6x. Check! So now we put all this information together and.... $(x + 3)^2$. Don't forget to check this work! $(x + 3)^2 = (x + 3)(x + 3) = x^2 + 3x + 3x + 9 = x^2 + 6x + 9$

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/factoring/special-factorizations-and-binomials/

Solving Quadratic Equations By Factoring

A quadratic eq. has two answers, both of which will correctly solve the equation; remember the FOIL method, and this will be much easier.

KEY POINTS

- A quadratic equation is a polynomial equation of the second degree. A general quadratic equation can be written in the form: $ax^2 + bx + c = 0$.
- It is a lot easier to solve a quadratic equation if you remember the FOIL method, and try and reverse this process.
- When you plug either of your answers back into the original equation, they both must work. This might seem like a lot of work right now, but after trying a few examples, it will be a lot easier to quickly recognize patterns and do it much more quickly. The key is to practice, practice, practice.

A **quadratic** equation is a polynomial equation of the second **degree**. A general quadratic equation can be written in the form:

$$ax^2 + bx + c = 0$$

where x represents a variable or an unknown, and a, b, and c are constants with a ≠ 0. (If a = 0, the equation is a linear equation.) Some plots of quadratic functions are shown in *Figure 1.17*.

The constants a, b, and c are called respectively, the quadratic coefficient, the linear coefficient and the constant term or free term. The term "quadratic" comes from quadratus, which is the Latin word for "square". Quadratic equations can be solved by **factoring**, completing the square, graphing, Newton's method, and using the quadratic formula.

To solve quadratic equations by factoring, lets remember the FOIL property. FOIL is an acronym for the order in which to multiply out a function. It stands for First, Outer, Inner, Last. Before we jump into factoring out a quadratic equation, lets try to FOIL a function:

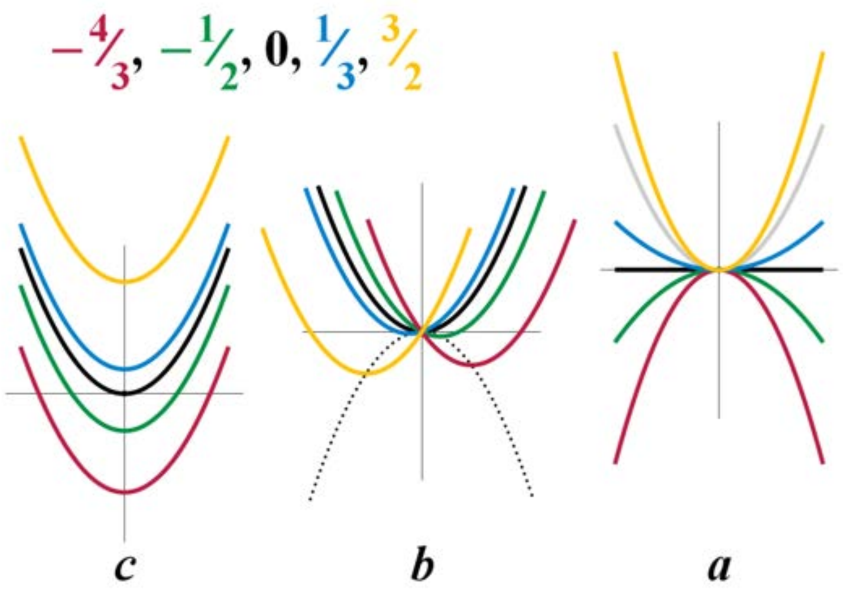
$$(x - 3)(x - 4) = 1x^2 - 4x - 3x + 12 = 1x^2 - 7x + 12 = 0$$

Remembering this, lets think about how we need to go about factoring out a quadratic equation. Lets do this by using the same example:

$$x^2 - 7x + 12 = 0$$

1. The first step is to set up the format: $(\quad)(\quad) = 0$
2. Now lets look at how to fill in the variable slots in this format. To do this, we need to take the square root of the first term,

Figure 1.17 Plots of Quadratic Equations



Plots of quadratic function $ax^2 + bx + c$, varying each coefficient separately.

or find two terms that will multiply together to the first term of our quadratic. In our example, it is simple, the square root of x2 is x, so lets input this into our format: $(x \quad)(x \quad) = 0$

3. Now, lets look at the constants to put in our equation. We need to have two numbers that (1) add up to the coefficient of our middle term and (2) multiple together to give us the same product of our last term. A good way to do this is make a list of all the numbers whose product is the last term, and make a note of what they add up to. In this example, our last term is 12, so lets make a list of all the numbers that multiply up to 12: 1*12; 2*6; 3* 4. We can easily see that the only option here that adds up to 7 is 3+4. So lets put this into our equation: $(x \quad 3)(x \quad 4) = 0$

4. The last thing to do is input the signs to our equation. A good thing to remember is that two negatives make a positive. So if our last term has a positive sign, then both of the signs in our factored equation are the same. If the last term is has a negative sign, the one of the equations will be negative, and one will be positive. Since we have a positive last term, we know that our equation is going to have the same signs in both terms. Since the middle term is negative, we know that both of our equations will have a negative sign. So now we can complete the factored equation: $(x - 3)(x - 4) = 0$

5. Now we really need to multiply this back out, and make sure it makes sense. ALWAYS CHECK YOUR WORK! When factored out, this returns the original equation, so we know it is correct.

6. Now, to solve the equation, set each term equal to zero: x-3=0 and x-4=0. So, x=3 and x=4

When you plug either of these terms back into the original equation, they both work. This might seem like a lot of work right now, but after trying a few examples, it will be a lot easier to quickly

recognize patterns and do it much more quickly. The key is to practice, practice, practice.

Section 5

Rational Expressions

Domain of a Rational Expression

Simplifying, Multiplying, and Dividing

Adding and Subtracting With Like and Unlike Denominators

Complex Rational Expressions

Solving Equations with Rational Expressions; Problems Involving Proportions

https://www.boundless.com/algebra/the-building-blocks-of-algebra/rational-expressions/

Domain of a Rational Expression

A rational expression cannot have a denominator of zero, so its domain is all numbers that do not cause the denominator to equal zero.

KEY POINTS

- A rational expression is the quotient of two polynomials. It can be expressed as.
- A rational expression's domain is set such that the denominator cannot equal zero. Therefore, $Q(x) \neq 0$.
- To determine the domain of a rational expression, set the denominator equal to zero, then solve for x. All values of x except for those that satisfy Q(x)=0 are the domain of the expression.

A **rational expression** is one which can be written as the ratio of two **polynomial** functions. Despite being called a rational expression, neither the coefficients of the polynomials nor the values taken by the function are necessarily rational numbers. In the case of one variable, x, an expression is called rational if and only if it can be written in the form

$$\frac{P(x)}{Q(x)}$$

where P(x) and Q(x) are polynomial functions in x and Q(x) is not the zero polynomial (Q(x)=0).

The domain of a rational expression of is the set of all points for which the denominator is not zero, where one assumes that the fraction is written in its lower degree terms, that is, P(x) and Q(x) have several factors of the positive degree. If the denominator of the equation becomes equal to zero, the

Examples

The rational expression

$$\frac{x^3 - 2x}{2(x^2 - 5)}$$

is not defined at x2=5, so it is not defined at

$$x = \pm\sqrt{5}$$

Therefore its domain is all numbers not equal to the square root of five or the negative square root of five.

The rational expression

$$\frac{x^2 - 2}{x}$$

is not defined at x=0, again because we would have to divide by 0. Therefore, the domain of this expression is all numbers not equal to zero.

For the expression

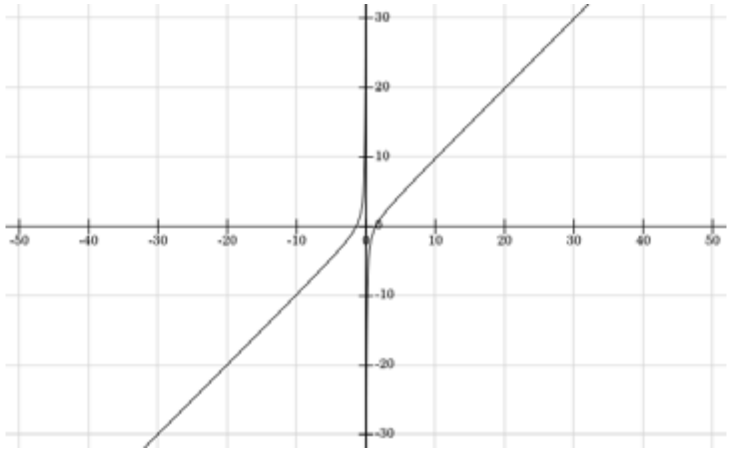
$$\frac{x^2 + 2}{x^2 + 1}$$

has a domain that includes all rational numbers. However, if x is equal to the square root of negative one, an irrational number, then the quotient is equal to zero. Therefore, the square root of negative one is outside the domain of this expression.

EXAMPLE

What is the domains of $\frac{x^2 - 2}{x}$? One way to determine this is to look at it graphically, and we can see that the graph is discontinuous at x=0, indicating that the domain is all numbers other than x=0. This makes sense, because at x=0 we would have to divide by zero, which is undefined.

Figure 1.18 $(x^2 - 2)/x$



To determine the domain of this function, we can graph it and look for where the function appears to go to infinity. Indeed, at x=0 the denominator will equal zero, and this is therefore outside of the domain of the function.

Simplifying, Multiplying, and Dividing

A rational expression can be treated like a fraction, and can be manipulated via multiplication and division.

KEY POINTS

- A rational expression is a quotient of two polynomials, of the form P(x)/Q(x).
- Rational expressions can often be simplified by removing terms that can be factored out of the numerator and denominator. These can either be numbers or functions of x.
- A rational expression can also be multiplied and divided, just like a normal fraction. When multiplying two rational expressions together, multiply the numerator of each together, then the denominator of each together. Sometimes, it is possible to simplify them after multiplying them together.

Just like a fraction involving numbers, a fraction involving polynomials (a rational expression) can be simplified, multiplied, and divided. The rules for performing these operations often mirror the rules for simplifying, multiplying, and dividing fractions, though instead, we are now factoring out polynomial expressions.

Simplifying a Rational Expression

As a first example, the rational expression

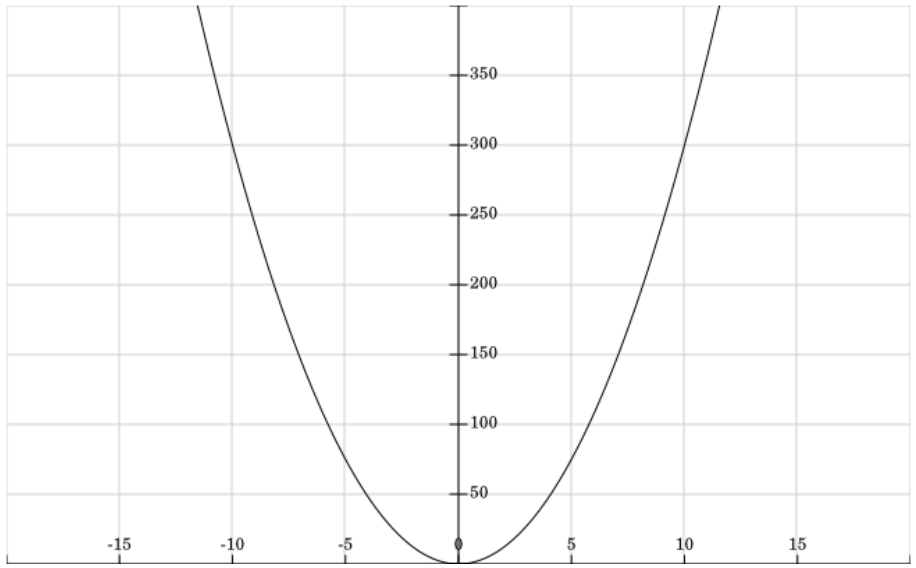
$$\frac{3x^3}{x}$$

can be simplified by canceling out one factor of x in the numerator and denominator, which gives the expression

$$3x^2.$$

The domain of the equation, however, does not include x, as this would cause division by 0 in the original equation. We can see that the latter form is a simplified version of the former graphically, as they appear the same (*Figure 1.19*).

Figure 1.19 Graph to Illustrate Simplification



Both $3x^2$ and $3x^3/x$

A more complicated example,

$$\frac{x^2 + 5x + 6}{2x^2 + 5x + 2}$$

must first be factored to provide the expression

$$\frac{(x+2)(x+3)}{(2x+1)(x+2)},$$

which, after canceling the common factor of (x+2) from both the numerator and denominator, gives the simplified expression

$$\frac{x+3}{2x+1}$$

which is a simplified form of the expression shown above.

Multiplying and Dividing Rational Expressions

Just like fractions, rational expressions can be multiplied and divided. First, we will multiply by whole numbers, then we will multiply one rational expression by another.

For example, the rational expression

$$\frac{x^2 + 3}{2x - 3}$$

can be multiplied by the fraction $\frac{2}{3}$ to provide

$\frac{2(x^2+3)}{3(2x-3)}$, which can be multiplied through to give $\frac{2x^2+6}{6x-9}$.

In this example, we multiplied the numerators together and the denominators together, but we did not multiply the numerator by the denominator or vice-versa.

If we want to multiply two rational expressions together, the rules are the same, but the operations are typically somewhat more complicated. As an example, we will look at the expression $\frac{x+1}{x-1}$ and multiply it by the expression $\frac{x+2}{x+3}$. The product of these two, initially, is $\frac{(x+1)(x+2)}{(x-1)(x+3)}$, which can be written out as $\frac{x^2+3x+2}{x^2+2x-3}$.

Notice that this expression cannot be simplified further.

Adding and Subtracting With Like and Unlike Denominators

Adding and subtracting rational expressions follows all of the same rules as adding and subtracting fractions.

KEY POINTS

- Always factor rational expressions before doing anything else.
- When two rational expressions are to be added or subtracted, they must be multiplied by a constant [in form f(x)/f(x)] so that they both have the same denominator.
- Once two rational expressions have the same denominator, the numerators can be added or subtracted together, leaving the denominator alone. Then simplify.

Adding and subtracting fractions should be a familiar process, and we will use this concept as a lead-in to start discussing the addition and subtraction of rational expressions.

$$\frac{1}{2}+\frac{1}{3}=\frac{3}{6}+\frac{2}{6}=\frac{5}{6}$$

The key is finding the least common denominator: the smallest multiple of both denominators. Then you rewrite the two fractions with this denominator. Finally, you add the fractions by adding the numerators and leaving the denominator alone.

But how do you find the least common denominator? Consider this problem:

$$\frac{5}{12}+\frac{7}{30}=?$$

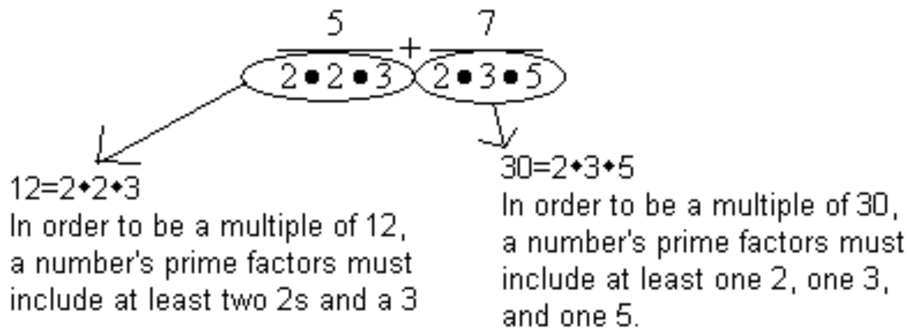
You could probably find the least common denominator if you played around with the numbers long enough. Here we will show you a systematic method for finding least common denominators—a method that works with rational expressions just as well as it does with numbers. We start, as usual, by factoring. For each of the denominators, we find all the **prime factors**, the prime numbers that multiply to give that number.

$$\frac{5}{2\cdot 2\cdot 3}+\frac{7}{2\cdot 3\cdot 5}$$

If you are not familiar with the concept of prime factors, it may take a few minutes to get used to. $2\times2\times3$ is 12, broken into its prime factors: that is, it is the list of prime numbers that multiply to give 12. Similarly, the prime factors of 30 are $2\times3\times5$.

Why does that help? Because 12=2×2×3, any number whose prime factors include two 2s and one 3 will be a multiple of 12. Similarly, any number whose prime factors include a 2, a 3, and a 5 will be a multiple of 30. See *Figure 1.20* for further details.

Figure 1.20 Prime Factors of Fractions



Finding the prime factors of the denominators of two fractions enables us to find a common denominator.

The least common denominator is the smallest number that meets both these criteria: it must have two 2s, one 3, and one 5. Hence, the least common denominator must be 2×2×3×5=60, and we can finish the problem like this.

$$\frac{5}{2\cdot2\cdot3}+\frac{7}{2\cdot3\cdot5}=\frac{5}{2\cdot2\cdot3}\cdot\frac{5}{5}+\frac{7}{2\cdot3\cdot5}\cdot\frac{2}{2}=\frac{25}{60}+\frac{14}{60}=\frac{39}{60}=\frac{13}{20}$$

This may look like a very strange way of solving problems that you have known how to solve since the third grade. However, you should spend a few minutes carefully following that solution, focusing on the question: why is 2×2×3×5 guaranteed to be the least common denominator? Because once you understand that, you have the key concept required to add and subtract rational expressions.

Addition and Subtraction of Rational Expressions

When applying this strategy to rational expressions, first look at the denominators of the two rational expressions and see if they are the same. If they are the same, then simply add or subtract the numerators from each other, leaving the denominator the same. If the two denominators are different, however, then a strategy similar to the one shown above is applicable.

For example if you wanted to perform the subtraction $\frac{3}{x^2+12x+36}-\frac{4x}{x^3+4x^2-12x}$ we begin rational expression problems of this type by factoring!

The denominators then become $\frac{3}{(x+6)^2}-\frac{4x}{x(x+6)(x-2)}$

The least common denominator must have two (x+6) s, one x, and one (x−2). We then rewrite both fractions with the common denominator, giving $\frac{3(x)(x-2)}{(x+6)^2(x)(x-2)}-\frac{4x(x+6)}{x(x+6)^2(x-2)}$

Subtracting fractions is easy when you have a common denominator! It is best to leave the bottom alone, since it is factored. The top, however, consists of two separate factored pieces,

and will be simpler if we multiply them out so we can combine them.

$$\frac{3(x)(x-2)-4x(x+6)}{x(x-2)(x+6)^2}$$

After multiplying out, we obtain $\frac{3x^2-6x-(4x^2+24x)}{x(x-2)(x+6)^2}$

A common student mistake here is forgetting the parentheses. The entire second term is subtracted; without the parentheses, the 24x ends up being added.

$\frac{-x^2-30x}{x(x-2)(x+6)^2}$ Almost done! But finally, we note that we can factor the top again. If we factor out an x it will cancel with the x in the denominator.

$$\frac{-x-30}{(x-2)(x+6)^2}$$

The problem is long, and the math is complicated. So after following all the steps, it is worth stepping back to realize that even this problem results simply from the two rules we started with.

First, always factor rational expressions before doing anything else.

Second, follow the regular processes for fractions: in this case, the procedure for subtracting fractions, which involves finding a common denominator. After that, you subtract the numerators while leaving the denominator alone, and then simplify.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/rational-expressions/adding-and-subtracting-with-like-and-unlike-denominators/

Complex Rational Expressions

A complex fraction is one where the numerator, denominator, or both are fractions. These fractions can contain variables, constants or both.

KEY POINTS

- A few examples of complex fractions, or complex rational expressions can be found below: $\frac{1-\frac{1}{x}}{1-\frac{1}{x^2}}$.
- Now, before we can solve these complex rational expressions, we first want to simplify them as much as possible. The best way to do this, is to get rid of the fractions in the numerator, denominator, or both if at all possible. The way this is done is by using simple algebraic techniques.
- One of the techniques that can be used is called the Combine-Divide Method.
  (1)Combine the terms in the numerator.
  (2)Combine the terms in the denominator.
  (3)Divide the numerator by the denominator.

We know from previous sections, that a simple fraction is in the form $\frac{P}{Q}$, where Q does not equal zero. The reason for this is simple, if the denominator were 0, the fraction would not a defined, meaning it would not be a real number.

A **complex fraction** is one in which the numerator, denominator, or both are fractions. These fractions can contain variables, constants or a mixture of both. A few examples of complex fractions, or complex rational expressions can be found below:

$$\frac{\frac{8}{15}}{\frac{2}{3}} \text{ and } \frac{1-\frac{1}{x}}{1-\frac{1}{x^2}}.$$

Now, before we can solve these complex rational expressions, we first want to simplify them as much as possible. The best way to do this, is to get rid of the fractions in the numerator, denominator, or both if at all possible. The way this is done is by using simple algebraic techniques.

One of the techniques that can be used is called the Combine-Divide Method.

1. Combine the terms in the numerator.
2. Combine the terms in the denominator.
3. Divide the numerator by the denominator.

Lets try to apply this to the first example complex fraction we saw:

$$\frac{\frac{8}{15}}{\frac{2}{3}}$$

Since there are no terms to combine in either the numerator or denominator, lets go right to step 3. Divide the numerator by the denominator: $\frac{8}{15} \div \frac{2}{3}$.

From previous sections, we know that dividing by a fraction is the same as multiplying by the inverse of that fraction, which would turn this expression into:

$$\frac{8}{15} * \frac{3}{2} = \frac{4}{5} * \frac{1}{1}$$

$$= \frac{4}{5}$$

Solving Equations with Rational Expressions; Problems Involving Proportions

Rational expressions, like proportions, are extremely useful applications of algebra, that can be solved using simple algebraic techniques.

KEY POINTS

- A rational equation means that you are setting two rational expressions equal to each other. Proportions are perfect examples of a rational expression. Even if they look different, they can be simplified down into the same expression: $(\frac{2}{4}) = (\frac{1}{2})$.
- If you have a rational equation where the denominators are the same, then the numerators must be the same. This in turn suggests a strategy: find a common denominator, and then set the numerators equal using algebraic techniques.
- Remember, all normal algebraic rules apply to solving rational equations. Such as, you still can not divide by 0.

A rational equation means that you are setting two **rational expressions** equal to each other. The goal is to solve for x; that is, find the value(s) that make the equation true.

Suppose you are told that:

$$\frac{x}{5} = \frac{3}{5}$$

If you think about it, the x in this equation has to be a 3. That is to say, if x=3 then this equation is true; for any other x value, this equation is false.If this is not so apparent to you, you can always solve it the old fashioned way, by working it out. Start by isolating the variable you are solving for: $x = (\frac{3}{5})3$ which simplifies down to

$$x = 3$$

This leads us to a very general rule: If you have a rational equation where the denominators are the same, then the numerators must be the same.

This in turn suggests a strategy: find a common denominator, and then set the numerators equal.

For example, consider the rational equation

$$\frac{3}{x^2 + 12x + 36} = \frac{4x}{x^3 + 4x^2 - 12x}$$

by factoring the denominators,we find that we must multiply the left side of the equation by $\frac{x(x-2)}{x(x-2)}$ and the right side of the equation by $\frac{x+6}{x+6}$, giving

$$\frac{3(x)(x-2)}{(x+6)^2(x)(x-2)} = \frac{4x(x+6)}{x(x+6)^2(x-2)}$$

Based on the rule above—since the denominators are equal, we can now assume the numerators are equal, so we know that $3(x)(x-2) = 4x(x+6)$ or, multiplied out, that $3x^2 - 6x = 4x^2 + 24x$

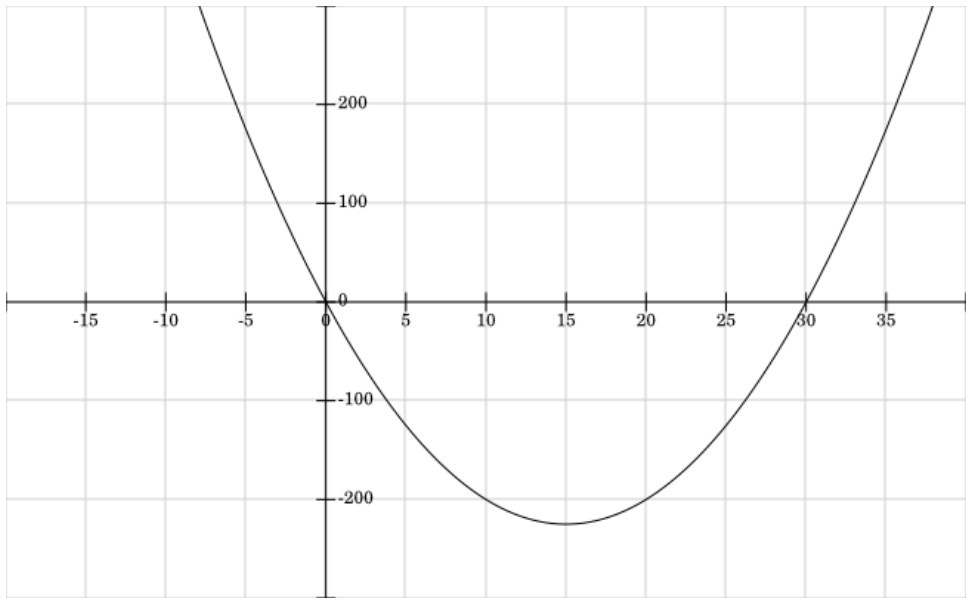
What we're dealing with, in this case, is a quadratic equation. As always, move everything to one side, giving $x^2 + 30x = 0$

and then factor. A common mistake in this kind of problem is to divide both sides by x; this loses one of the two solutions.

$$x(x-30) = 0$$

Two solutions to the quadratic equation. However, in this case, x=0 is not valid, since it was not in the domain of the original right-hand fraction. (Why?) So this problem actually has only one solution, x=–30.This is shown in *Figure 1.21*.

Figure 1.21 Graphical determination of the solutions of x(x-30)=0



To determine the solutions to the equation x(x-30)=0, we can graph it and look for where the dependent variable crosses the x-axis. We find this is true at 0 and 30.

As always, it is vital to remember what we have found here. We started with the equation $\frac{3(x)(x-2)}{(x+6)^2(x)(x-2)} = \frac{4x(x+6)}{x(x+6)^2(x-2)}$. We have concluded now that if you plug x=−30 into that equation, you will get a true equation (you can verify this on your calculator). For any other value, this equation will evaluate false.

EXAMPLE

When given the rational equation: $(\frac{a}{b}) = (\frac{c}{d})$ This can be solved by either finding a common denominator, or by setting it up like: $ad = cb$ and then solving it algebraically.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/rational-expressions/solving-equations-with-rational-expressions-problems-involving-proportions/

Section 6

Radical Notation and Exponents

Simplifying Expressions

Adding, Subtracting, and Multiplication

Solving Problems with Radicals

Rationalizing Denominators or Numerators

Rational Exponents

Complex Numbers

Radical Functions

https://www.boundless.com/algebra/the-building-blocks-of-algebra/radical-notation-and-exponents/

Simplifying Expressions

Both radicals and exponents can be simplified to a basic expression by understanding the rules of working with each.

KEY POINTS

- A simplified radical adheres by the following: There is no factor of the radicand that can be written as a power greater than or equal to the index, there are no fractions under the radical sign, and there are no radicals in the denominator.
- Multiplication and division of exponents of the same base can be simplified using addition and subtraction.
- Raising a power to a power is simplified by muliplication.

In mathematics, the nth root of a number x is a number r that, when raised to the power of n, equals x:

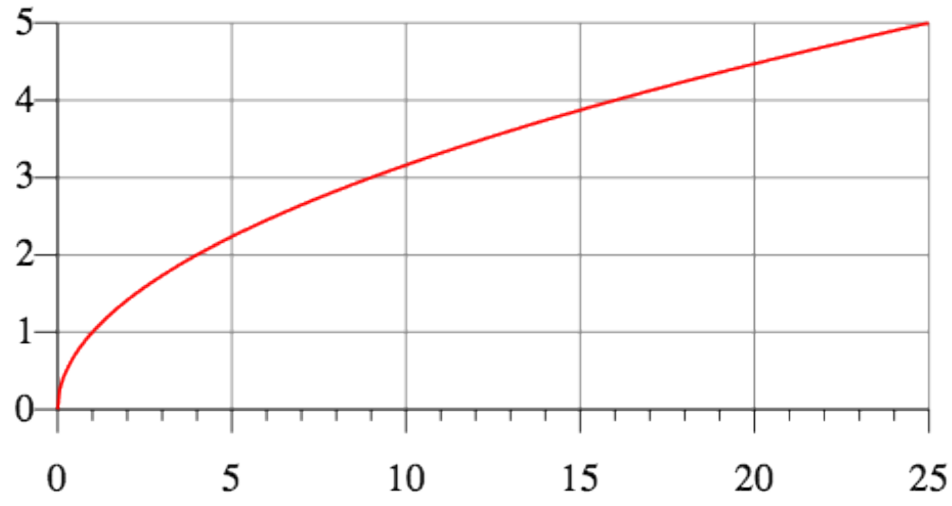
$r^n = x$

where n is the degree of the root. A root of degree 2 is called a square root (*Figure 1.22*) and a root of degree 3, a cube root. Roots of higher degrees are referred to using ordinal numbers, as in fourth root, twentieth root, etc.

A radical expression is said to be in simplified form if:

1. There is no factor of the **radicand** that can be written as a power greater than or equal to the index.
2. There are no fractions under the radical sign.
3. There are no radicals in the denominator.

Figure 1.22 Square root



The graph of the function, made up of half a parabola with a vertical directrix.

For example, to write the radical expression $\sqrt{\frac{32}{5}}$ in simplified form, we can proceed as follows. First, look for a perfect square under the square root sign and remove it:

$$\sqrt{\frac{32}{5}} = \sqrt{\frac{16 \cdot 2}{5}} = 4\sqrt{\frac{2}{5}}$$

Next, there is a fraction under the radical sign, which we change as follows:

$$4\sqrt{\frac{2}{5}} = \frac{4\sqrt{2}}{\sqrt{5}}$$

Finally, we remove the radical from the denominator as follows:

$$4\sqrt{\frac{2}{5}} = \frac{4\sqrt{2}}{\sqrt{5}} \cdot \frac{\sqrt{5}}{\sqrt{5}} = \frac{4\sqrt{10}}{5}$$

For simplifying exponents, just follow the rules of exponents and get down to the most basic expression.

Multiplying Powers with the Same Base

$a^m \cdot a^n = a^{m+n}$

a^m means that you have a factor of m times. If you add n more factors of a then you have n+m factors of a.

Dividing Powers with the Same Base

$$\frac{a^m}{a^n} = a^{m-n}$$

In the same way that $a^m \cdot a^n = a^{m+n}$, because you are adding on factors of a, dividing is taking away factors of a. If you have n factors of a in the denominator, then you can cross out n factors from the numerator. If there were m factors in the numerator, now you have m-n factors in the numerator.

Raising a Power to a Power

$(a^n)^m = a^{n \cdot m}$

If you think about an exponent as telling you that you have so many factors of the base, then $(a^n)^m$ means that you have factors m of a^n. So you have m groups of a^n and each one of those has n groups of a. So you have m groups of n groups of a. So you have $n \cdot m$ groups of a, or $a^{n \cdot m}$.

Products Raised to Powers

$(ab)^n = a^n b^n$

You can multiply numbers in any order you please. Instead of multiplying together n factors equal to ab, you could multiply all of the a s together, multiply all the b s together, then finish by multiplying a^n times b^n.

Adding, Subtracting, and Multiplication

Radicals and exponents have particular requirements for addition and subtraction while multiplication is carried out more freely.

KEY POINTS

- To add radicals, the radicand (the number that is under the radical) must be the same for each radical.
- Subtraction follows the same rules as addition: the radicand must be the same.
- Multiplication of radicals simply requires that we multiply the term under the radical signs.
- When multiplying exponents, the bases must be the same then we just add together the exponents. am means that you have a factor of m times. If you add n more factors of a then you have n+m factors of a.

Roots are the inverse operation for exponents. An expression with roots is called a **radical** expression. It's easy, although perhaps tedious, to compute exponents given a root. For instance 7*7*7*7 = 49*49 = 2401. So, we know the fourth root of 2401 is 7, and the square root of 2401 is 49. What is the third root of 2401? Finding the value for a particular root is difficult. This is because exponentiation is a different kind of function than addition, subtraction, multiplication, and division.

Let's go through some basic mathematical operations with radicals and exponents.

Addition

To add radicals, the radicand (the number that is under the radical) must be the same for each radical, so, a generic equation will look like *Figure 1.23*.

Figure 1.23 Addition of Radicals

$$a\sqrt{b} + c\sqrt{b} = a + c\sqrt{b}$$

The generic equation for adding radicals

Figure 1.24 shows some examples of radical manipulations.

Let's plug some numbers in place of the variables:

$\sqrt{3} + 2\sqrt{3} = 3\sqrt{3}$

Figure 1.24 Examples of additions of radicals

a. $7\sqrt{11}+6\sqrt{11}=13\sqrt{11}$

b. $5\sqrt{2}-7\sqrt{2}=-2\sqrt{2}$

c. $\sqrt[3]{5}-4x\sqrt[3]{5}+8\sqrt[3]{5}=(1-4x+8)\sqrt[3]{5}=(9-4x)\sqrt[3]{5}$

d. $8\sqrt[6]{5x}-5\sqrt[6]{5x}+4\sqrt[3]{5x}=3\sqrt[6]{5x}+4\sqrt[3]{5x}$

Examples of additions of radicals, subtractions of radicals, and combinations of the two.

Subtraction

Subtraction follows the same rules as addition:

$$a\sqrt{b}-c\sqrt{b}=(a-c)\sqrt{b}$$

For example:

$$3\sqrt{3}-2\sqrt{3}=\sqrt{3}$$

Multiplication

Multiplication of radicals simply requires that we multiply the variable under the radical signs.

$$\sqrt{a}\cdot\sqrt{b}=\sqrt{a\cdot b}$$

When multiplying exponents, the bases must be the same then we just add together the exponents. a^m means that you have a factor of m times. If you add n more factors of a then you have n+m factors of a.

$$a^m\cdot a^n=a^{m+n}$$

Some examples with real numbers:

$$\sqrt{3}\cdot\sqrt{6}=\sqrt{18}$$

This equation can actually be simplified further; we will go over simplification in another section.

Let's throw in some variables to demonstrate that the concept carries over:

$$\sqrt{2x}\cdot\sqrt{3x^2}=\sqrt{6x^3}$$

Here we incorporated what we have learned about radicals and exponents; multiply the terms that are under the radical, in the case of the exponents, which have a the same base, x, we add them together.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/radical-notation-and-exponents/adding-subtracting-and-multiplication/

Solving Problems with Radicals

Roots are written using a radical sign, and a number denoting which root to solve for. When none is given, it is an implied square root.

KEY POINTS

- Roots are usually written using the radical symbol, but can also be written by raising the number to a fraction. Then, the root is the inverse of the raised power. Like this: $\sqrt{x} = x^{\frac{1}{2}}$.
- To solve an equation with a radical: isolate the radical on one side of the equation, get rid of your radical, solve the remaining equation.
- To eliminate a square root, square the radical, to eliminate a cubed root, cube the radical - don't forget to do the exact same thing to the other side of the equation!

Roots are written using a **radical** sign. If there is no denotation, it is implied that you are finding the square root. Otherwise, a number will appear denoting which root to solve for. Any expression containing a radical is called a radical expression.

The best way to solve an equation, is to start by simplifying it as much as possible. You want to start by getting rid of the radical. Do

this by treating the radical as if it where a variable. Isolate it on one side and go from there.

Let's look at how to do it step-by-step:

1. Isolate the radical on one side of the equation.
2. Get rid of your radical (some of the rules listed below may help in this).
3. Repeat steps 1&2 if you have another radical.
4. Solve the remaining equation.
5. Double check equation by plugging in your answer.

And remember, always treat each side of the equation the same, here's some helpful reminders for general equation solving: *Figure 1.25*.

Some helpful rules:

1. $\sqrt{x} \cdot \sqrt{x} = (\sqrt{x})^2 = x$
2. $\sqrt{\frac{x}{y}} = \frac{\sqrt{x}}{\sqrt{y}}$
3. $x^{\frac{m}{n}} = (\sqrt[n]{x})^m$
4. $\sqrt{x}\sqrt{y} = \sqrt{xy}$
5. $\sqrt[m]{\sqrt[n]{x}} = \sqrt[m \cdot n]{x} = x^{\frac{1}{m \cdot n}}$

Let's run through an example:

Solve the following: $\sqrt{2x + 5} = 7$

1. Isolate the radical (already done).
2. Get rid of the radical: $(\sqrt{2x + 5})^2 = (7)^2$

$2x + 5 = 49$

No more radicals? Great, solve for x:

$2x + 5 - 5 = 49 - 5$

$2x = 44$

$\frac{2x}{2} = \frac{44}{2}$

$x = 22$

Figure 1.25 Properties of Equality

Property Name	Definition	Example
Reflexive	a = a	8=8
Symmetric	If a = b, then b = a	If (3)(2) = 6, then 6 = (3)(2)
Transitive	If a = b & b = c, then a = c	If 8 = (4)(2) and (4)(2) = (2)(4), then 8 = (2)(4)
Substitution	If a = b, then one can replace *a* with *b* or vice versa	If a = b and 1 + a = 3, then 1 + b = 3
Addition	You can add one number to both sides of the equation.	$x-6=14$ $x-6+6=14+6$ $x=20$
Subtraction	You can subtract one number from both sides of the equation.	$x+6=14$ $x+6-6=14-6$ $x=8$
Multiplication	You can multiply both sides of the equation by a number.	$\frac{x}{6}=18$ $6(\frac{x}{6})=(18)(6)$ $x=108$
Division	You can divide both sides of the equation by a number.	$6x=18$ $6x(\frac{1}{6})=(\frac{18}{6})$ $x=3$

Make sure to use these properties when solving an equation.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/radical-notation-and-exponents/solving-problems-with-radicals/

Rationalizing Denominators or Numerators

Root rationalization is a process by which roots in the denominator of an irrational fraction are eliminated.

KEY POINTS

- Both the numerator and the denominator can be rationalized but it is more common to rationalize the denominator.
- To rationalize the denominator, multiply both the numerator and denominator by the radical in the denominator.
- You may not to be to rationalize both the numerator and the denominator.

Obviously, with a graphing calculator, we can solve almost any expression, but it is always good to be able to do these things by hand before we start just plugging things in. Calculators are great, but if you plug something in wrong, you won't be able to recognize that the answer is incorrect if you aren't able to do it by hand as well. It is good to practice these so that you can start to recognize patterns and understand what it is you are doing as opposed to just trusting a device.

A lot of times in mathematics, we are given terms in the form of fractions that have radicals in the **denominator**. When we are given expressions like this, before we start to solve it, it is easiest to write them in the most simplified way possible. You want to start by taking the radicals OUT of the denominator. This can be done using simple, basic, algebraic techniques. We know that what ever we do to one side of an algebraic equation, we must also do to the other side. This same principal can be applied to fractions. What ever we do to the **numerator**, we must also do to the denominator, and visa versa. As many things are, this is easier to show than the explain. Lets start with the following example:

You are given this fraction, $\frac{10}{\sqrt{a}} * \frac{\sqrt{a}}{\sqrt{a}} = \frac{\sqrt{a} * 10}{\sqrt{a}^2} = \frac{10 * \sqrt{a}}{a}$

A more complicated example of denominator **rationalization** is shown in *Figure 1.26*.

Figure 1.26 Rationalization of a Denominator

For a denominator that is:

$$\sqrt{2} + \sqrt{3}$$

Rationalisation can be achieved by multiplying by the *Conjugate*:

$$\sqrt{2} - \sqrt{3}$$

and applying the difference of two squares identity, which here will yield −1. To get this result, the entire fraction should be multiplied by

$$\frac{\sqrt{2} - \sqrt{3}}{\sqrt{2} - \sqrt{3}} = 1.$$

This technique works much more generally. It can easily be adapted to remove one square root at a time, i.e. to rationalise

$$x + \sqrt{y}$$

by multiplication by

$$x - \sqrt{y}$$

Example:

$$\frac{3}{\sqrt{3} + \sqrt{5}}$$

The fraction must be multiplied by a quotient containing $\sqrt{3} - \sqrt{5}$.

$$\frac{3}{\sqrt{3} + \sqrt{5}} \cdot \frac{\sqrt{3} - \sqrt{5}}{\sqrt{3} - \sqrt{5}} = \frac{3(\sqrt{3} - \sqrt{5})}{\sqrt{3}^2 - \sqrt{5}^2}$$

Now, we can proceed to remove the square roots in the denominator:

$$\frac{3(\sqrt{3} - \sqrt{5})}{\sqrt{3}^2 - \sqrt{5}^2} = \frac{3(\sqrt{3} - \sqrt{5})}{3 - 5} = \frac{3(\sqrt{3} - \sqrt{5})}{-2}$$

A more complicated example of denominator rationalization

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/radical-notation-and-exponents/rationalizing-denominators-or-numerators/

Rational Exponents

Exponents are shorthand for repeated multiplication and can be used to express an n-th root of a number: b is a number x such that x^n = b.

KEY POINTS

- The number in larger font is called the base. The number in superscript (that is, the smaller number written above) is called the exponent.
- If b is a positive real number and n is a positive integer, then there is exactly one positive real solution to $x^n = b$. This solution is called the principal n-th root of b. It is denoted n√b, where √ is the radical symbol; alternatively, it may be written $b^{1/n}$.
- A power of a positive real number b with a rational exponent m/n in lowest terms satisfies $b^{\frac{m}{n}} = (b^m)^{\frac{1}{n}} = \sqrt[n]{b^m}$.

Exponents are a shorthand used for repeated multiplication. Remember that when you were first introduced to multiplication it was as a shorthand for repeated addition. For example, you learned that: 4 × 5 = 5 + 5 + 5 + 5. The expression "4 × " told us how many times we needed to add. Exponents are the same type of shorthand for multiplication. Exponents are written in superscript after a regular-sized number.

For example: 2^3 = 2 x 2 x 2. The number in larger font is called the base. The number in superscript (that is, the smaller number written above) is called the exponent. The exponent tells us how many times the base is multiplied by itself. In this example, 2 is the base and 3 is the exponent.The expression 2^3 is read aloud as "2 raised to the third power", or simply "2 cubed".

Here are some other examples: 6 × 6 = 6^2 (This would read aloud as "six times six is six raised to the second power" or more simply "six times six is six squared".) 7 × 7 × 7 × 7 = 7^4 (This would read aloud as "seven times seven times seven times seven equals seven raised to the fourth power". There are no alternate expression for raised to the fourth power. It is only the second and third powers that usually get abbreviated because they come up more often. When it is clear what is being talked about, people often drop the words "raised" and "power" and might simply say "seven to the fourth".)

Rational Exponents

A rational exponent is a rational number that can be used as another way to write roots. An n-th root of a number b is a number x such that $x^n = b$.

If b is a positive real number and n is a positive integer, then there is exactly one positive real solution to $x^n = b$. This solution is called the principal n-th root of b. It is denoted n√b, where √ is the

radical symbol; alternatively, it may be written $b^{1/n}$. For example: $4^{1/2} = 2$, $8^{1/3} = 2$.

When one speaks of the n-th root of a positive real number b, one usually means the principal n-th root.

- If n is even, then $x^n = b$ has two real solutions;
- If b is positive, which are the positive and negative nth roots.
- The equation has no solution in real numbers if b is negative.
- If n is odd, then $x^n = b$ has one real solution.
- The solution is positive if b is positive and negative if b is negative.

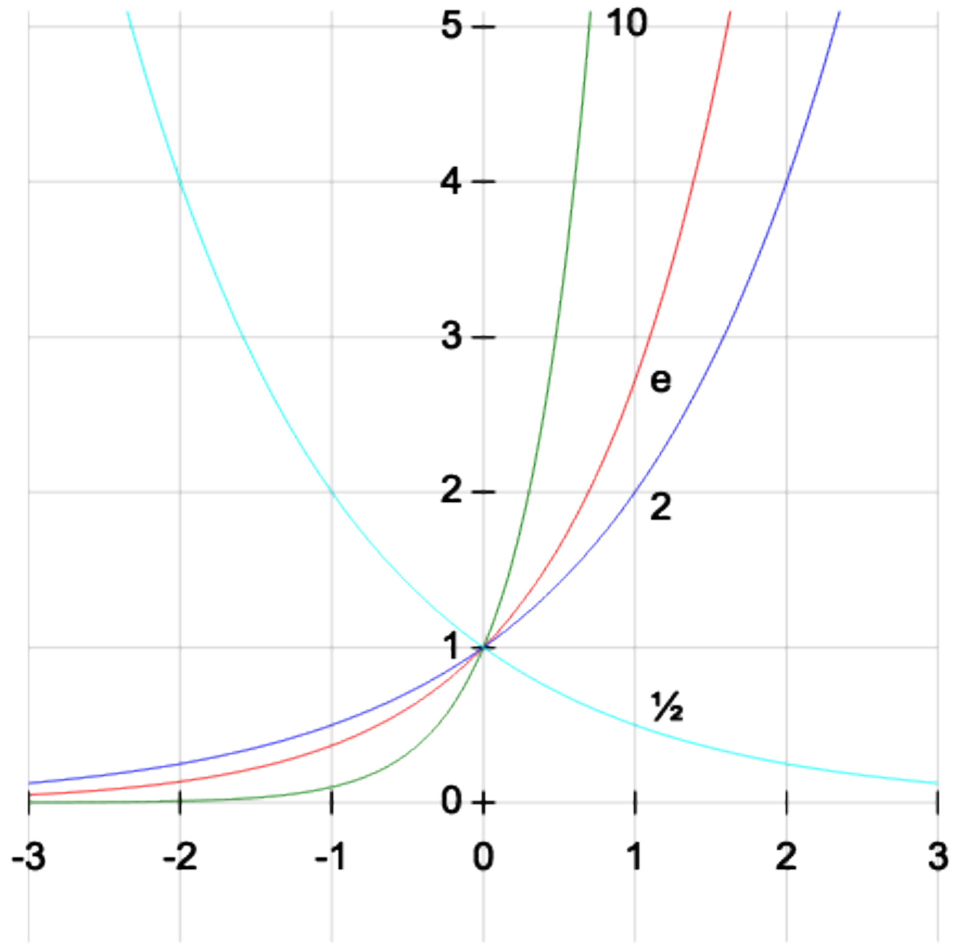
Rational powers m/n, where m/n is in lowest terms, are positive if m is even, negative for negative b if m and n are odd, and can be either sign if b is positive and n is even. $(-27)^{1/3} = -3$, $(-27)^{2/3} = 9$, and $4^{3/2}$ has two roots 8 and −8. Since there is no real number x such that $x^2 = -1$, the definition of $b^{m/n}$ when b is negative and n is even must use the imaginary unit i. A power of a positive real number b with a rational exponent m/n in lowest terms satisfies

$$b^{\frac{m}{n}} = (b^m)^{\frac{1}{n}} = \sqrt[n]{b^m}$$

where m is an integer and n is a positive integer.

Examples of exponents graphed can be seen in this figure *Figure 1.27*.

Figure 1.27 Exponential Graph



Graphs of $y = b^x$ for various bases b: base 10 (green), base e (red), base 2 (blue), and base ½ (cyan). Each curve passes through the point (0, 1) because any nonzero number raised to the power of 0 is 1. At x=1, the y-value equals the base because any number raised to the power of 1 is the number itself.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/radical-notation-and-exponents/rational-exponents/

Complex Numbers

We say that all numbers of the form a +bi, where a and b are any real numbers, is the set of complex numbers.

KEY POINTS

- $x^2 = -1$ results in a "number" that would not be a real number, referred to as an imaginary number.
- When working with imaginary numbers, both -i and i have equal standing: $\sqrt{-1} = \pm i$.
- Complex numbers can be added, subtracted, multiplied, and divided (except by 0).

Complex numbers are the extension of the real numbers, i.e., the number line, into a number plane. They allow us to turn the rules of plane geometry into arithmetic. Complex numbers have fundamental importance in describing the laws of the universe at the subatomic level, including the propagation of light and quantum mechanics. They also have practical uses in many fields, including signal processing and electrical engineering.

Currently, one is able to solve many different kinds of equations for x, such as $x + 7 = 12$ or $2^x = 4$. In each of these cases the solution for is a real number. However, there is no real number x that satisfies the equation $x^2 = -1$, since the square of any real number is non-

negative. Conceptually, it would be nice to have some kind of number to be the solution of $x^2 = -1$. This "number" would not be a real number, however, and is referred to as an imaginary number. Then the real number system is extended to accommodate this special number. It turns out that there will be two imaginary solutions of the equation. One of them will be called i, following the normal rules for arithmetic, the other solution is -i. One may be inclined to say that $\sqrt{-1} = i$. However, that would be incorrect, because in words this says that "the square root of -1 is i", but there is no basis for preferring i over -i (or vice versa) as the square root of -1. Rather, the two square roots have equal standing.

All numbers of the form a +bi, where a and b are any real numbers, are a set of complex numbers, and denoted as set C. The real numbers R may be considered to be the subset of complex numbers C = {a +bi} for which b = 0. Complex numbers can be added, subtracted, multiplied, and divided (except by 0). For a negative root like $\sqrt{-4}$, split the number into two parts such that one part is $\sqrt{-1}$ like $\sqrt{-4} = \sqrt{4 \cdot -1} = \sqrt{4} \cdot \sqrt{-1}$ which leads to $2i$. More examples of complex numbers can be seen in *Figure 1.28*.

Figure 1.28 Complex Numbers

- 1 + 4i: a complex number, real part 1, imaginary part 4
- 2 - 2i: a complex number, real part 2, imaginary part -2.
- -4i: a complex number, real part 0, imaginary part -4
- 2: a complex number (also a real number), real part 2, imaginary part 0.

Notice that the number 2 is a complex number **and** a real number. This fact is clearer if we write 2 = 2 + 0i.

Examples of Complex Numbers

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/radical-notation-and-exponents/complex-numbers/

Radical Functions

An expression with roots is called a radical function, there are many kinds of roots, square root and cube root being the most common.

KEY POINTS

- Roots are the inverse operation for exponents. If $r^n = x$.
- If the square root of a number is taken, the result is a number which when squared gives the first number.
- The cube root is the number which, when cubed, or multiplied by itself and then multiplied by itself again, gives back the original number.
- If a root of a whole number is squared root, which is not itself the square of a rational number, the answer will have an infinite number of decimal places.. Such a number is described as irrational.

Roots are the inverse operation for exponents. An expression with roots is called a **radical** expression. It's easy, although perhaps tedious, to compute exponents given a root. For instance $7 \cdot 7 \cdot 7 \cdot 7 = 49 \cdot 49 = 2401$.

If fourth root of 2401 is 7, and the square root of 2401 is 49, then what is the third root of 2401?

Finding the value for a particular root is difficult. This is because exponentiation is a different kind of function than addition, subtraction, multiplication, and division. When graphing functions, expressions that use exponentiation use curves instead of lines. Using algebra will show that not all of these expressions are functions and that knowing when an expression is a relation or a function allows certain types of assumptions to be made. These assumptions can be used to build mental models for topics that would otherwise be impossible to understand.

For now, deal with roots by turning them back into exponents. If a root is defined as the nth root of X, it is represented as $\sqrt[n]{x} = r$. Get rid of the root by raising the answer to the nth power, i.e. $r^n = x$.

Square root

If the square root of a number is taken, the result is a number which when squared gives the first number. This can be written symbolically as: $\sqrt{x} = y$ if $y^2 = x$.

In the series of real numbers $y^2 \geq 0$, regardless of the value of y. As such, when $x < 0$ then $\sqrt{x}$ cannot be defined.

Such examples of square roots can be seen in *Figure 1.29*.

Figure 1.29 Square Root

Examples:

- $\sqrt{9} = 3$ since $3^2 = 3 \cdot 3 = 9$.
- If $x = 25$ then $\sqrt{x - 16} = \sqrt{25 - 16} = \sqrt{9} = 3$
- If $x = 7$ then $\sqrt{x - 16}$ is undefined because $\sqrt{x - 16} = \sqrt{7 - 16} = \sqrt{-9}$, but there is no number y so that $y^2 = -9$. Notice that the answer is **not** -3 since $-3 \cdot -3 = 9$ and not -9

Examples of square roots

Cube roots

Roots do not have to be square. The cube root of a number ($\sqrt[3]{\ }$) can also be taken. The cube root is the number which, when cubed, or multiplied by itself and then multiplied by itself again, gives back the original number. For example, the cube root of 8 is 2 because $2 \cdot 2 \cdot 2 = 8$, or $\sqrt[3]{8} = 2$.

Other roots

There are an infinite number of possible roots all in the form of $\sqrt[n]{a}$ which corresponds to $a^{\frac{1}{n}}$, when expressed using exponents. If $\sqrt[n]{a} = b$ then $b^n = a$. The only exception is 0. $\sqrt[0]{a}$ is undefined, as it corresponds to $a^{\frac{1}{0}}$, resulting in a division by zero. Even if attempting to discover the 0th root of 1, no progress will be made, as practically any number to the power of zero equals 1, leaving only an undefined result.

Irrational numbers

If a root of a whole number is squared root, which is not itself the square of a rational number, the answer will have an infinite number of decimal places. Such a number is described as irrational and is defined as a number which cannot be written as a rational number: $\frac{a}{b}$, where a and b are integers. However, using a calculator can approximate the square root of a non-square number: $\sqrt{3} = 1.73205080757$

The result of taking the square root is written with the approximately equal sign because the result is an irrational value which cannot be written in decimal notation exactly. Writing the square root of 3 or any other non-square number as $\sqrt{3}$ is the simplest way to represent the exact value. Irrational numbers also appear when attempting to take cube roots or other roots. However, they are not restricted to roots, and may also appear in other mathematical constants (e.g. π, e, φ, etc.).

Basics of Equation Solving

Linear and Quadratic Equations

Fractions

Simplifying Algebraic Expressions

Additive and Multiplicative Properties of Equality

Percent and Mixture Problem Solving

https://www.boundless.com/algebra/the-building-blocks-of-algebra/basics-of-equation-solving/

Linear and Quadratic Equations

Equations often express relationships between given quantities; two types of equations are linear and quadratic.

KEY POINTS

- A linear equation is an algebraic equation in which each term is either a constant or the product of a constant and (the first power of) a single variable. Linear equations can have one or more variables.
- Using the laws of elementary algebra, linear equations can be rewritten into several different forms, such as general, slope-intercept, and point-slope form.
- A quadratic equation is a univariate polynomial equation of the second degree. The constants a, b, and c are respectively called the quadratic coefficient, the linear coefficient, and the constant term (or free term).

Equations often express relationships between given quantities (the knowns) and quantities yet to be determined (the unknowns). By convention, unknowns are denoted by letters at the end of the alphabet, x, y, z, w, …, while knowns are denoted by letters at the beginning, a, b, c, d, …. The process of expressing the unknowns in terms of the knowns is called solving the equation. In an equation with a single unknown, a value of that unknown for which the equation is true is called a solution or root of the equation. In a set simultaneous equations, or system of equations, multiple equations are given with multiple unknowns. A solution to the system is an assignment of values to all the unknowns so that all of the equations are true. Two kinds of equations are linear and quadratic.

Linear Equations

A **linear equation** is an algebraic equation in which each term is either a constant or the product of a constant and (the first power of) a single variable. Linear equations can have one or more variables. Linear equations do not include exponents.

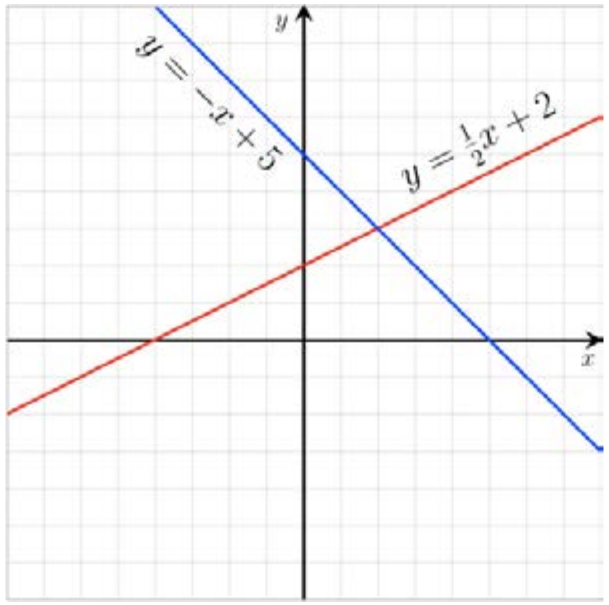
A common form of a linear equation in the two variables x and y is:

$$y = mx + b$$

where m and b designate constants. The origin of the name "linear" comes from the fact that the set of solutions of such an equation forms a straight line in the plane. In this particular equation, the constant m determines the slope or gradient of that line, and the constant term b determines the point at which the line crosses the y-axis, otherwise known as the y-intercept. Since terms of linear equations cannot contain products of distinct or equal variables, nor any power (other than 1) or other function of a variable, equations

involving terms such as xy, x^2, $y^{1/3}$, and sin(x) are nonlinear. An example of a graphed linear equation is presented in *Figure 1.30*.

Figure 1.30 Linear Function Graph



Graph sample of linear equations.

Using the laws of elementary algebra, linear equations can be rewritten into several different forms. A few of these forms are below. These equations are often referred to as the "equations of the straight line." In what follows, x, y, t, and θ are variables, m is the slope, and b is the y-intercept.

General (or Standard) Form

$$Ax + By = C$$

where A and B are both not equal to zero. The equation is usually written so that A ≥ 0, by convention. If A is nonzero, then the x-intercept, or the x-coordinate of the point where the graph crosses the x-axis (where y is zero), is C/A. If B is nonzero, then the y-intercept, or the y-coordinate of the point where the graph crosses the y-axis (where x is zero), is C/B, and the slope of the line is –A/B.

Slope–Intercept Form

$$y = mx + b$$

This can be seen by letting x = 0, which immediately gives y = b. Vertical lines, having undefined slopes, cannot be represented by this form.

Point–Slope Form

$$y - y_1 = m(x - x_1)$$

where (x_1, y_1) is any point on the line. The point-slope form expresses the fact that the difference in the y-coordinate between two points on a line (that is, y – y_1) is proportional to the difference in the x-coordinate (that is, x – x_1).

Quadratic Equations

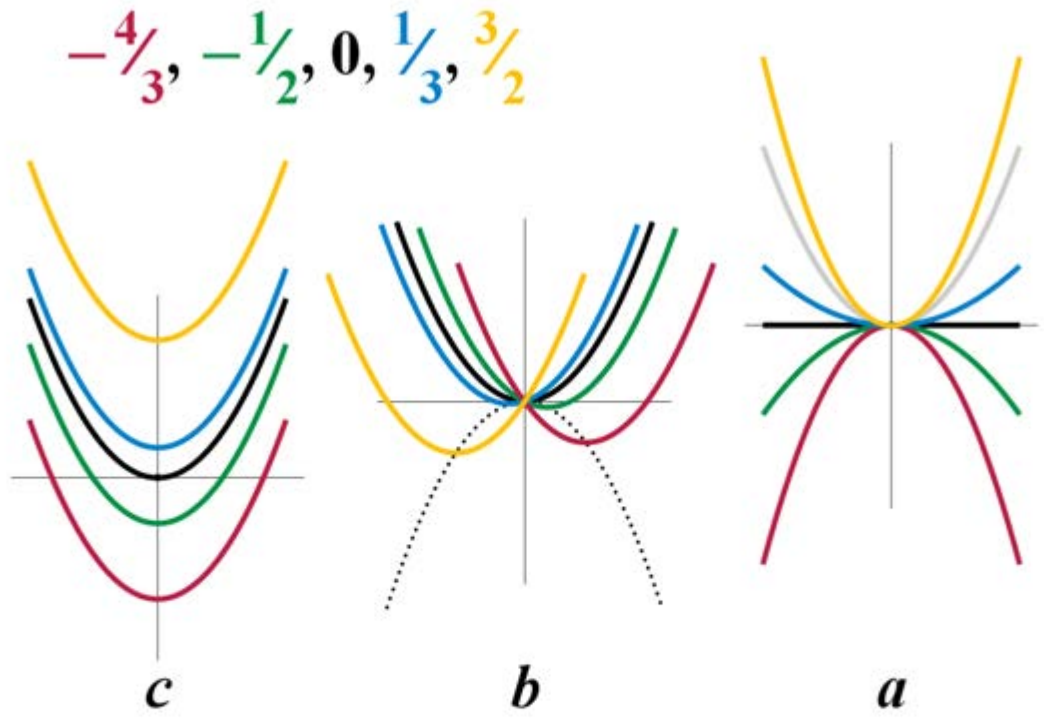
A **quadratic equation** is a **univariate** polynomial equation of the second degree. A general quadratic equation can be written in the form:

$$ax^2 + bx + c = 0$$

where x represents a variable or an unknown, and a, b, and c are constants with a ≠ 0. (If a = 0, the equation is a linear equation.) The constants a, b, and c are respectively called the quadratic

coefficient, the linear coefficient, and the constant term (or free term). The term "quadratic" comes from quadratus, which is Latin for "square." Quadratic equations can be solved by factoring, completing the square, graphing, Newton's method, and using the quadratic formula. Examples of graphed quadratic equations can be seen in *Figure 1.31*.

Figure 1.31 Plots of Quadratic Equations



Plots of the real-valued quadratic function $ax^2 + bx + c$, varying each coefficient separately.

A quadratic equation with real or complex coefficients has two solutions, called roots. These two solutions may or may not be distinct, and they may or may not be real. Having

$$ax^2 + bx + c = 0$$

the roots are given by the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where the symbol "±" indicates that both

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

$$x = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

are solutions of the quadratic equation.

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/basics-of-equation-solving/linear-and-quadratic-equations/

Fractions

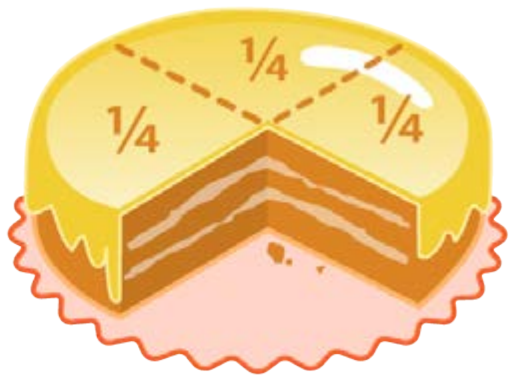
A fraction represents a part of a whole and consists of an integer numerator and non-zero integer denominator.

KEY POINTS

- Addition and subtraction require like quantities, a common denominator, in order to add or subtract fractions containing unlike quantities (e.g. quarters and thirds), it is necessary to convert all amounts to like quantities.
- Multiplication requires multiplying the numerators and then the denominators. A shortcut is to use cancellation by reducing to the lowest number prior to multiplication.
- To divide a fraction by a whole number, either divide the numerator by the number, if it goes evenly into the numerator, or multiply the denominator by the number.

A **fraction** represents a part of a whole or, more generally, any number of equal parts. A common or vulgar fraction, such as 1/2, 8/5, 3/4, consists of an integer numerator and a non-zero integer denominator—the numerator representing a number of equal parts and the denominator indicating how many of those parts make up a whole. An example can be seen in *Figure 1.32*, where a cake is divided into quarters.

Figure 1.32 Quarters of a Cake



A cake with one fourth removed. The remaining three fourths are shown. Dotted lines indicate where the cake may be cut in order to divide it into equal parts. Each fourth of the cake is denoted by the fraction ¼.

The set of all numbers which can be expressed in the form a/b, where a and b are integers and b is not zero, is called the set of rational numbers and is represented by the symbol Q, which stands for quotient. The test for whether a number is a rational is that it can be written in that form (i.e., as a common fraction). However, the word fraction is also used to describe mathematical expressions that are not rational numbers, for example algebraic fractions, and expressions that contain irrational numbers, such as $\sqrt{2}/2$ and $\pi/4$.

Addition

The first rule of addition is that only like quantities can be added; for example, various quantities of quarters. Unlike quantities, such as adding thirds to quarters, must first be converted to like quantities as described below: Imagine a pocket containing two quarters, and another pocket containing three quarters; in total,

there are five quarters. Since four quarters is equivalent to one (dollar), this can be represented as follows:

$$\frac{2}{4}+\frac{3}{4}=\frac{5}{4}=1\frac{1}{4}$$

To add fractions containing unlike quantities (e.g. quarters and thirds), it is necessary to convert all amounts to like quantities. It is easy to work out the chosen type of fraction to which to convert. Simply multiply together the two denominators, bottom number, of each fraction. For adding quarters to thirds, both types of fraction are converted to twelfths, thus:

$$\frac{1}{3}+\frac{1}{4}=\frac{1*4}{3*4}+\frac{1*3}{4*3}=\frac{4}{12}+\frac{3}{12}=\frac{7}{12}$$

This method can be expressed algebraically:

$$\frac{a}{b}+\frac{c}{d}=\frac{ad+cd}{bd}$$

This method always works, but sometimes there is a smaller denominator, or a least common denominator, that can be used. For example, to add $\frac{3}{4}$ and $\frac{5}{12}$ the denominator 48 can be used (the product of 4 and 12), but the smaller denominator 12 may also be used, being the least common multiple of 4 and 12.

Subtraction

The process for subtracting fractions is, in essence, the same as that of adding them: find a common denominator, and change each fraction to an equivalent fraction with the chosen common denominator. The resulting fraction will have that denominator, and its numerator will be the result of subtracting the numerators of the original fractions. For instance,

$$\frac{2}{3}-\frac{1}{2}=\frac{2*2}{3*2}-\frac{1*3}{2*3}=\frac{4}{6}-\frac{3}{6}=\frac{1}{6}$$

Multiplication

To multiply fractions, multiply the numerators and multiply the denominators. Thus:

$$\frac{2}{3}\times\frac{3}{4}=\frac{6}{12}$$

A short cut for multiplying fractions is called "cancellation". In effect, one reduces the answer to lowest terms during multiplication. For example:

$$\frac{2}{3}\times\frac{3}{4}=\frac{1}{1}\times\frac{1}{2}=\frac{1}{2}$$

A two is a common factor in both the numerator of the left fraction and the denominator of the right and is divided out of both. Three is

a common factor of the left denominator and right numerator and is divided out of both. To multiply a fraction by a **whole number**, place the whole number over one and multiply.This method works because the fraction 6/1 means six equal parts, each one of which is a whole. When multiplying mixed numbers, it's best to convert the mixed number into an improper fraction.

Division

To divide a fraction by a whole number, either divide the numerator by the number, if it goes evenly into the numerator, or multiply the denominator by the number. For example,

$$\frac{10}{3} \div 5 = \frac{2}{3}$$

which also equals: $\frac{10}{3 * 5} = \frac{10}{15} = \frac{2}{3}$

To divide a number by a fraction, multiply that number by the reciprocal of that fraction. Thus,

$$\frac{1}{2} \div \frac{3}{4} = \frac{1}{2} \times \frac{4}{3} = \frac{4}{6} = \frac{2}{3}.$$

Simplifying Algebraic Expressions

Algebraic expressions may be simplified, based on the basic properties of arithmetic operations.

> KEY POINTS
>
> - Added terms are simplified using coefficients: $2x^2 + 3ab - x^2 + ab = x^2 + 4ab$.
> - Multiplied terms are simplified using exponents: $x \times x \times x = x^3$.
> - Using distributive law, brackets can be multiplied out: $6x^5 + 3x^2 = 3x^2(2x^3 + 1)$.

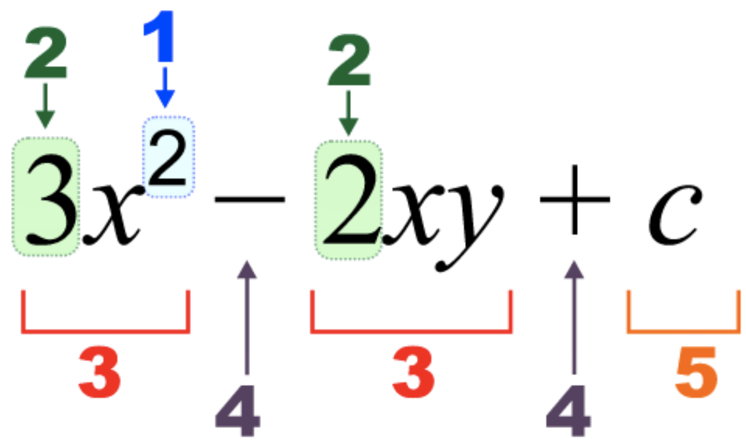


Figure 1.33
Algebraic Notation
1 – Exponent (power), 2 – Coefficient, 3 – term, 4 – operator, 5 – constant, x,y – variables

Algebraic notation follows certain rules and conventions, and has its own terminology. For example, look at *Figure 1.33*, as you can see, the expression consists of an exponent, coefficients, terms, operators, constants and variables.

A coefficient is a numerical value which multiplies a variable (the operator is omitted). A term is an addend or a summand, a group of coefficients, variables, constants and exponents that may be separated from the other terms by the plus and minus operators. Letters represent variables and constants. By convention, letters at the beginning of the alphabet are typically used to represent constants, and those toward the end of the alphabet are used to represent variables. They are usually written in italics.

Algebraic operations work in the same way as arithmetic operations, such as addition, subtraction, multiplication, division and exponentiation and are applied to algebraic variables and terms. Multiplication symbols are usually omitted, and implied when there is no space between two variables or terms, or when a coefficient is used. Usually terms with the highest power (exponent), are written on the left. When a coefficient is one, it is usually omitted. Likewise when the exponent (power) is one. When the exponent is zero, the result is always 1. However, being undefined, should not appear in an expression, and care should be

taken in simplifying expressions in which variables may appear in exponents.

Now that we understand each of the components of the expression, let's look at how we simplify them. Algebraic expressions may be evaluated and simplified, based on the basic properties of arithmetic operations (addition, subtraction, multiplication, division and exponentiation). Let's look at each individually:

- Added terms are simplified using coefficients. For example, $x + x + x$ can be simplified as $3x$ (where 3 is the coefficient).
- Multiplied terms are simplified using exponents. For example, $x \times x \times x$ is represented as x^3.
- Like terms are added together. For example, $2x^2 + 3ab - x^2 + ab$ is written as: $x^2 + 4ab$, because the terms containing are added together, and, the terms containing are added together.
- Brackets can be "multiplied out", using distributivity. For example, $x(2x + 3)$ can be written as $(x \times 2x) + (x \times 3)$ which can be written as: $2x^2 + 3x$
- Expressions can be factored. For example, $6x^5 + 3x^2$, by dividing both terms by $3x^2$ can be written as: $3x^2(2x^3 + 1)$.

Additive and Multiplicative Properties of Equality

The additive and multiplicative properties of equalities are common ways used to solve equations.

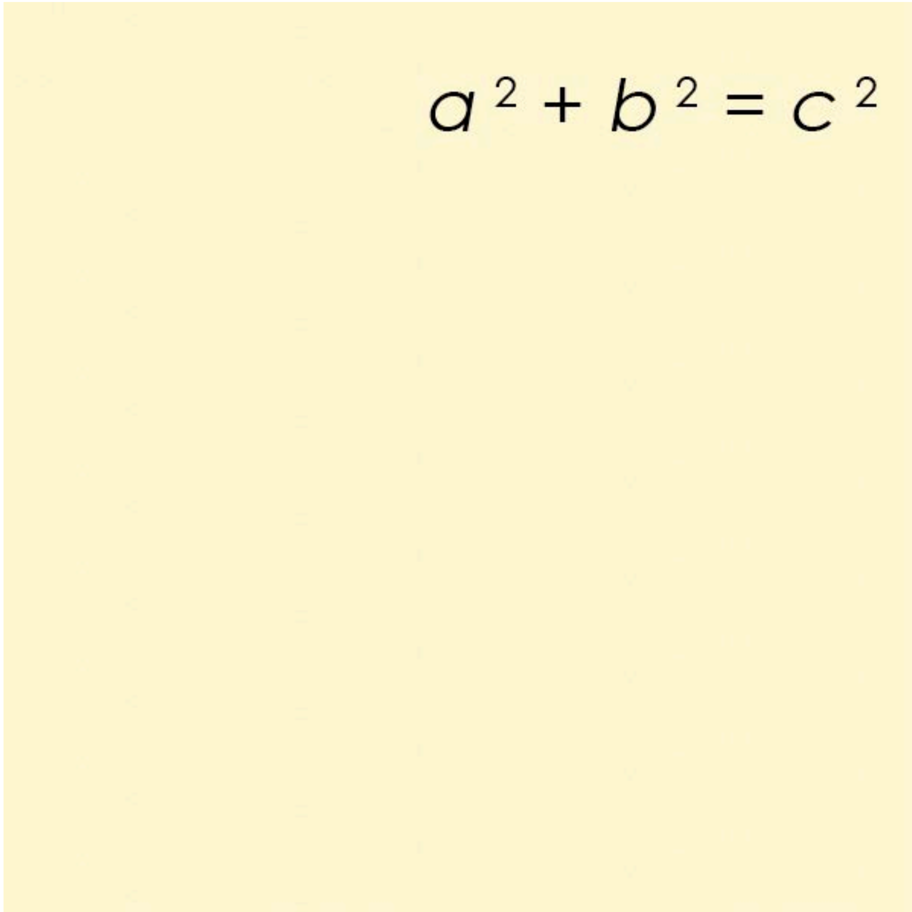
KEY POINTS

- Equations often express relationships between given quantities, the knowns, and quantities yet to be determined, the unknowns.
- The Additive Property of Equality states that any real number can be added to both sides.
- The Multiplicative Property of Equality states that any real number can be multiplied to both sides.

An equation, in a mathematical context, is generally understood to mean a mathematical statement that asserts the **equality** of two expressions. In modern notation, this is written by placing the expressions on either side of an equals sign (=), for example $x + 3 = 5$ asserts that x+3 is equal to 5. The = symbol was invented by Robert Recorde (1510–1558), who considered that nothing could be more equal than parallel straight lines with the same length.

Equations often express relationships between given quantities, the knowns, and quantities yet to be determined, the unknowns (*Figure 1.34*). By convention, unknowns are denoted by letters at the end of the alphabet, x, y, z, w, ..., while knowns are denoted by letters at the beginning, a, b, c, d, ... The process of expressing the unknowns in terms of the knowns is called solving the equation. In an equation

Figure 1.34 Equation Solving



Animation illustrating Pythagoras' rule for a right-angle triangle, which shows the algebraic relationship between the triangle's hypotenuse, and the other two sides.

with a single unknown, a value of that unknown for which the equation is true is called a solution or root of the equation. In a set of simultaneous equations, or system of equations, multiple equations are given with multiple unknowns. A solution to the system is an assignment of values to all the unknowns so that all of the equations are true.

If an equation in algebra is known to be true, the following operations may be used to produce another true equation:

- Any real number can be added to both sides (the Additive Property of Equality).
- Any real number can be subtracted from both sides.
- Any real number can be multiplied to both sides (the Multiplicative Property of Equality).
- Any non-zero real number can divide both sides.

The Additive Property

If $a = b$ then $a + c = b + c$

The Multiplicative Property

If $a = b$ then $a \cdot c = b \cdot c$

Percent and Mixture Problem Solving

Percentages are used to express how large/small one quantity is relative to another quantity.

KEY POINTS

- Mixture problems may have a mix of two or more things requiring you to find the percentage of one of those things.
- The first quantity of a percentage usually represents a part of, or a change in, the second quantity, which should be greater than zero.The percent value is computed by multiplying the numeric value of the ratio by 100.
- Due to inconsistent usage, it is not always clear from the context what a percentage is relative to. When speaking of a "10% rise" or a "10% fall" in a quantity, the usual interpretation is that this is relative to the initial value of that quantity.

Percentages

In mathematics, a percentage is a number or ratio as a fraction of 100. It is often denoted using the **percent** sign (*Figure 1.35*) or the abbreviation "pct." For example, 45% (read as "forty-five percent") is equal to 45/100, or 0.45. A related system which expresses a number as a fraction of 1,000 uses the terms "per mil" and "millage."

Figure 1.35
Percent Sign
The sign for percent (which means of 100).

Percentages are used to express how large/small one quantity is relative to another quantity. The first quantity usually represents a part of or a change in the second quantity, which should be greater than zero. For example, an increase of \$0.15 on a price of \$2.50 is an increase by a fraction of 0.15/2.50 = 0.06. Expressed as a percentage, this is therefore a 6% increase. Although percentages are usually used to express numbers between zero and one, any ratio can be expressed as a percentage. For instance, 111% is 1.11 and −0.35% is −0.0035. Although this is technically inaccurate as per the definition of percent, an alternative wording in terms of a change in an observed value is "an increase/decrease by a factor of..."

Mixture Problems

Mixture problems may have a mix of two or more things requiring you to find the percent of one of those things. The percent value is computed by multiplying the numeric value of the ratio by 100. For example, to find the percentage of 50 apples out of 1,250 apples, first compute the ratio 50/1250 =.04, and then multiply by 100 to obtain 4%. The percent value can also be found by multiplying first,

so in this example the 50 would be multiplied by 100 to give 5,000, and this result would be divided by 1,250 to give 4%.

To calculate a percentage of a percentage, convert both percentages to fractions of 100, or to decimals, and multiply them. For example, 50% of 40% is:(50/100) × (40/100) = 0.50 × 0.40 = 0.20 = 20/100 = 20%. It is not correct to divide by 100 and use the percent sign at the same time. (E.g. 25% = 25/100 = 0.25, not 25% / 100, which actually is (25/100) / 100 = 0.0025. A term such as (100/100)% would also be incorrect, as this would be read as (1) percent even if the intent was to say 100%.)

The easy way to calculate addition in percentage (discount 10% + 5%): For example, if a department store has a "10% + 5% discount," the total discount is not 15%. However, whenever we talk about a percentage, it is important to specify what it is relative to, i.e. what is the total that corresponds to 100%.

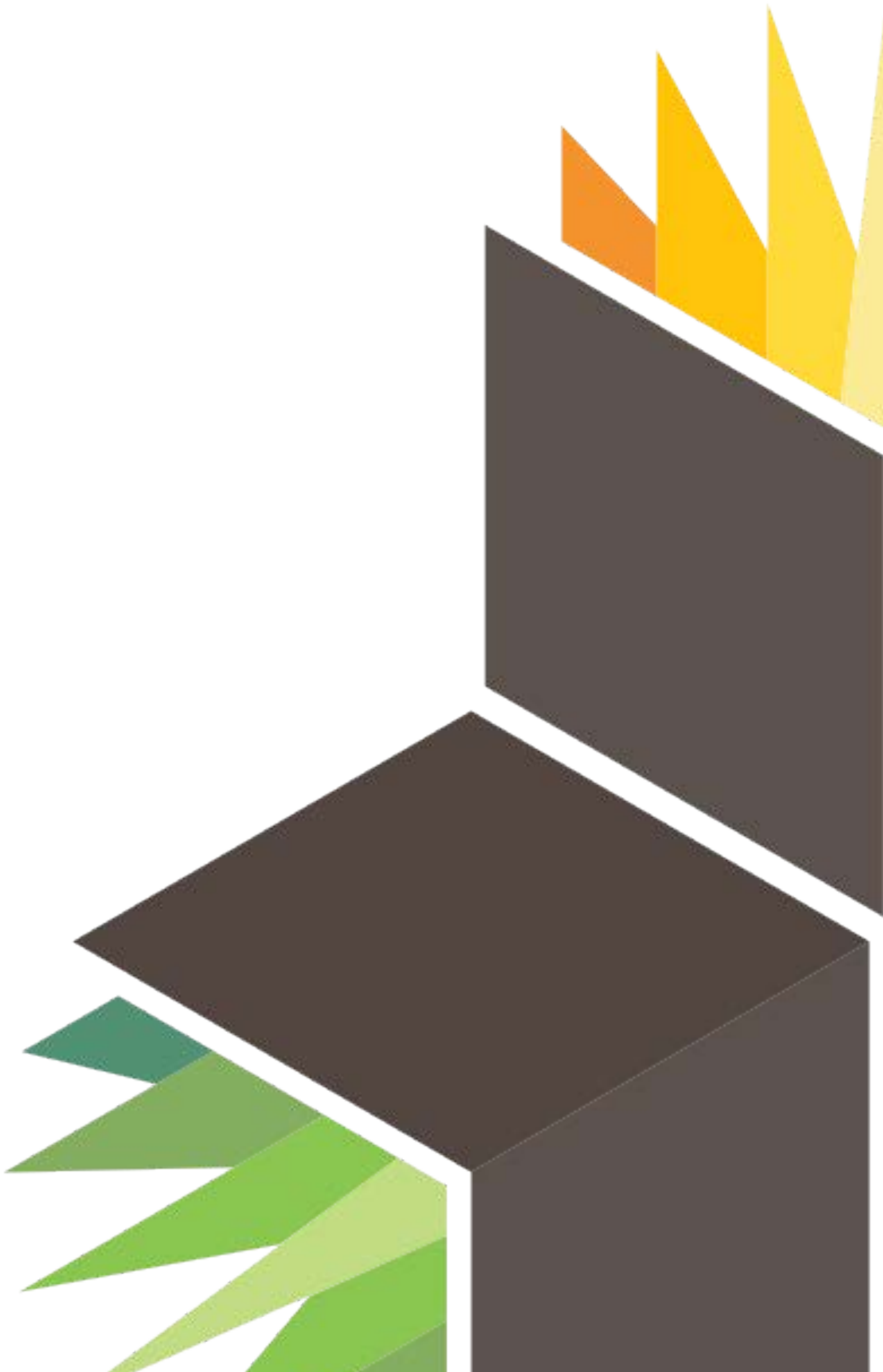
The following problem illustrates this point: In a certain college 60% of all students are female, and 10% of all students are computer science majors. If 5% of female students are computer science majors, what percentage of computer science majors are female? We are asked to compute the ratio of female computer science majors to all computer science majors. We know that 60% of all students are female, and among these 5% are computer science majors, so we conclude that (60/100) × (5/100) = 3/100 or 3% of all students are female computer science majors. Dividing this by the 10% of all students that are computer science majors, we arrive at the answer: 3%/10% = 30/100 or 30% of all computer science majors are female. This example is closely related to the concept of conditional probability.

Due to inconsistent usage, it is not always clear what a percentage is relative to. When speaking of a "10% rise" or a "10% fall" in a quantity, this is usually relative to the initial value of that quantity. For example, if an item is initially priced at $200 and the price rises 10% (an increase of $20), the new price will be $220. Note that this final price is 110% of the initial price (100% + 10% = 110%).

Source: https://www.boundless.com/algebra/the-building-blocks-of-algebra/basics-of-equation-solving/percent-and-mixture-problem-solving/

Chapter 2

Graphs, Functions, and Models



https://www.boundless.com/algebra/graphs-functions-and-models/

Section 1

Introduction to Graphing

https://www.boundless.com/algebra/graphs-functions-and-models/introduction-to-graphing/

The Cartesian System

The Cartesian coordinate system is used to specify points on a graph by showing their absolute distances from two axes.

KEY POINTS

- The Cartesian coordinate system is a 2-dimensional plane with a horizontal axis, known as the x-axis, and a vertical axis, known as the y-axis.
- A Cartesian coordinate system specifies each point uniquely in a plane with a pair of numerical coordinates, which are the signed distances from the point to the two axes.
- Points are represented by an ordered pair (x, y), where the x-coordinate is the point's distance from the y axis, and the y-coordinate is the distance from the x-axis.

Named for "the father of analytic geometry," 17th-century French mathematician René Descartes, the Cartesian coordinate system is a 2-dimensional plane with a horizontal axis and a vertical axis used for graphing. Both axes extend to infinity, but in graphs only segments of them are drawn, and sometimes arrows are used to indicate the infinite length. The horizontal axis is known as the **x-axis** and the vertical axis is known as the **y-axis**. The point where the axes intersect is known as the origin.

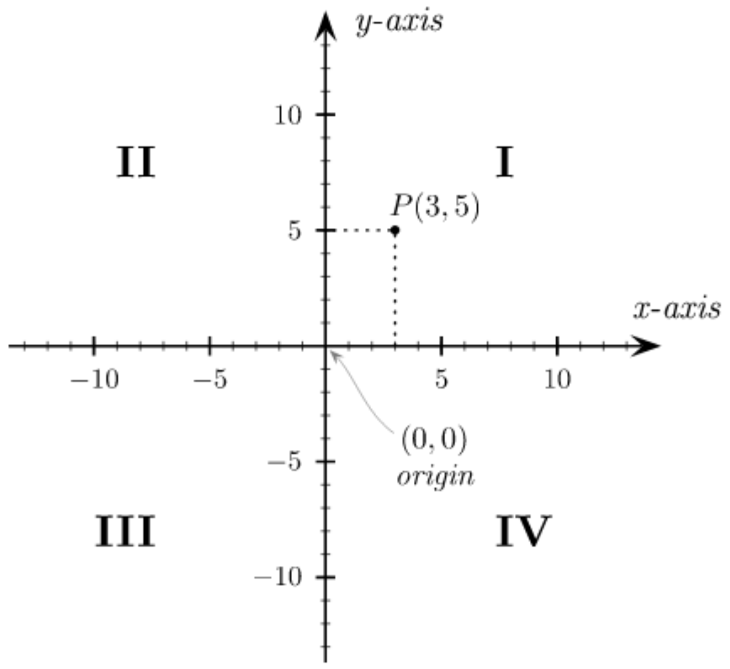
Cartesian coordinates are the foundation of analytic geometry and provide enlightening geometric interpretations for many other branches of mathematics, such as linear algebra, complex analysis, differential geometry, multivariate calculus, group theory, and more. A familiar example is the graph of a function, which you will learn in subsequent chapters. Cartesian coordinates are also essential tools for most applied disciplines that deal with geometry, including astronomy, physics, engineering, and many more. They are the most common coordinate system used in computer graphics, computer-aided geometric design, and other geometry-related data processing.

Coordinates of a Point

A Cartesian coordinate system specifies each point uniquely in a plane by a pair of numerical coordinates, which are the signed distances from the point to the two axes. Each point can be represented by an **ordered pair** (x, y), where the x-coordinate is the point's distance from the y axis, and the y-coordinate is the distance from the x-axis. The origin where the two axes meet is thus (0, 0). When the coordinates are integer numbers, they can be easily found on a graph by looking at the numbers on the axes. The ordered pair (3, 5) represents a point three units to the right of the origin and five units upwards from it. On the x-axis, numbers increase toward the right and decrease toward the left. On the y-

axis, numbers increase going upward and decrease going downward.

Figure 2.1 2D Cartesian Coordinates



The four quadrants of a Cartesian coordinate system. The arrows on the axes indicate that they extend forever in their respective directions (i.e. infinitely).

The intersection of the two axes splits the coordinate system into four **quadrants**, typically labelled I, II, III, and IV, starting from the upper right and continuing counter-clockwise, as seen in *Figure 2.1*. Any point in the first quadrant has both positive x and y coordinates. Points in the second quadrant have negative x and positive y coordinates. The third quadrant has both negative x and y coordinates, and the fourth quadrant contains points with positive x and negative y coordinates. Thus we know that the point (3, 5) is in the first quadrant.

Equations and their Solutions

Solutions to equations can be plotted on graphs for an easy visualization of how the function behaves.

KEY POINTS

- To solve an equation is to find what values (numbers, functions, sets, etc.) fulfill a condition stated in the form of an equation (two expressions related by equality).
- In mathematics, the graph of a function f is the collection of all ordered pairs (x, f(x)).
- Once a function has been graphed, solutions to any particular x or y value are readily available by looking at the graph.

In mathematics, to solve an **equation** is to find what values (numbers, functions, sets, etc.) fulfill a condition stated in the form of an equation (two **expressions** related by equality). These expressions contain one or more unknowns, which are free variables for which values are sought that cause the condition to be fulfilled. To be precise, what is sought are often not necessarily actual values, but, more in general, mathematical expressions, or a visualization of the solution.

Using Graphs

In mathematics, the **graph** of a function f is the collection of all ordered pairs (x, f(x)). In particular, if x is a real number, graph means the graphical representation of this collection, in the form of a curve on a Cartesian plane, together with Cartesian axes, etc. Graphing on a Cartesian plane is sometimes referred to as curve sketching.

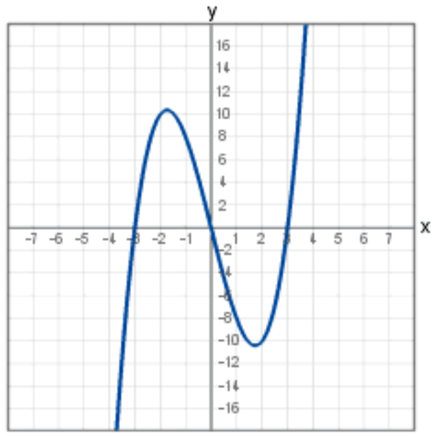


Figure 2.2 Graph of a Function

This is the graph of the function $f(x) = x^3 - 9x$

Say the graph of a function is {(1, a), (2, d), (3, c)}. You can then plot these points on a Cartesian plot as a graph. Now you know if you ever get the value x=1, you can look at the graph and see that the value is a, x=2 and you get d. What if x = 4? In this situation, there is no value on the graph for which x = 4, and so there is no solution at this point.

The graph of the cubic polynomial on the real line $f(x) = x^3 - 9x$ is {(x, x3-9x) : x is a real number). So how does this help us? With this information we can plot specific points by choosing an x value, finding the corresponding y value, and graphing it. The set of

infinite points is the curve that defines the function. If this set is plotted on a Cartesian plane, the result is *Figure 2.2*. Once we have a graph, we can then look at the graph to see how the function changes as x changes, as well as choose any x value and find its y value along the graph.

Graphing Equations

Equations can lead to very interesting graphs, such as circles, ellipses, parabolas, hyperbolas, sine waves, and much more!

KEY POINTS

- Graphs are important tools to visualizing equations.
- To graph an equation, choose a value for either x or y, solve for the variable you didn't choose to get a point, plot the point, and repeat until you have enough points to draw the graph by connecting the points.
- Graphs can be closed, like a circle, or open, like a parabola.

Now that we know what equations are, how do we go about visualizing them? With **graphs**! For an equation with two variables, x and y, we need a graph with two axes, an x axis and a y axis. The x-axis will be a horizontal line, the y-axis a vertical line, and where the two cross is the origin. Let's look at the equation $y = 2x$. To begin graphing this equation, all you need to do is choose an x or y value, plug it in, solve for the unknown variable, and plot the **point** on the graph. For instance, if you choose x = 1, then y = 2, and we plot the point (1, 2). If x = 0, y = 0, and we can plot this point as well (0, 0). We could keep doing this with a lot of points,

and if we used an infinite amount of points, we would get a line that passes through all the points. Obviously we can't plot an infinite amount of points, but if we plot just enough to see the overall shape of the graph of the equation, then we can "connect the dots" and visualize the equation.

A line is pretty boring, so let's look at some other equations. We won't go into detail as to why these equations are what they are, as that will be covered in subsequent sections. First let's look at the equation

$x^2 + y^2 = 100.$

Before we tell you what this type of equation actually is, let's choose some points and plot them. Easiest would be to choose x = 0, thus we're left with $y^2 = 100$, and $\sqrt{y^2} = \sqrt{100}$, and finally $y = \pm 10$ So we plot (0, 10) and (0, -10). Now let's choose another easy point, y = 0. By the same calculation as above we get the ordered pairs (10, 0) and (-10, 0). Plot these as well. We still don't have enough points to really see what's going on, so let's choose some more. How about x = 6. Solve for y: $36 + y^2 = 100$, then $y^2 = 64$, and finally $y = \pm 8$. So that's two new points (6, 8) and (6, -8). We get similar results with x = -6, to get (-6, 8) and (-6, -8). You may begin to see where this is going. With a few more points you can show that (8, 6), (8, -6), (-8, 6) and (-8, -6) are also all part of the equation. Now you can begin seeing that what we're drawing is a circle with radius 10 (*Figure 2.3*).

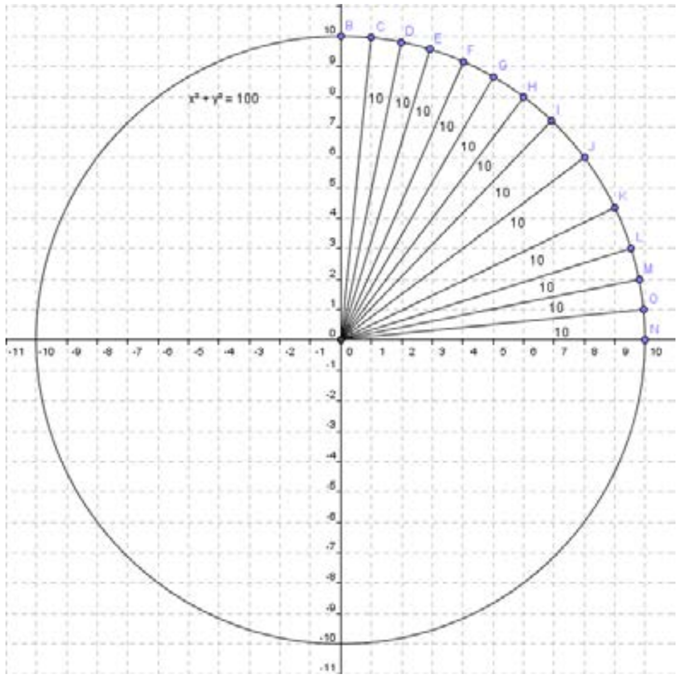


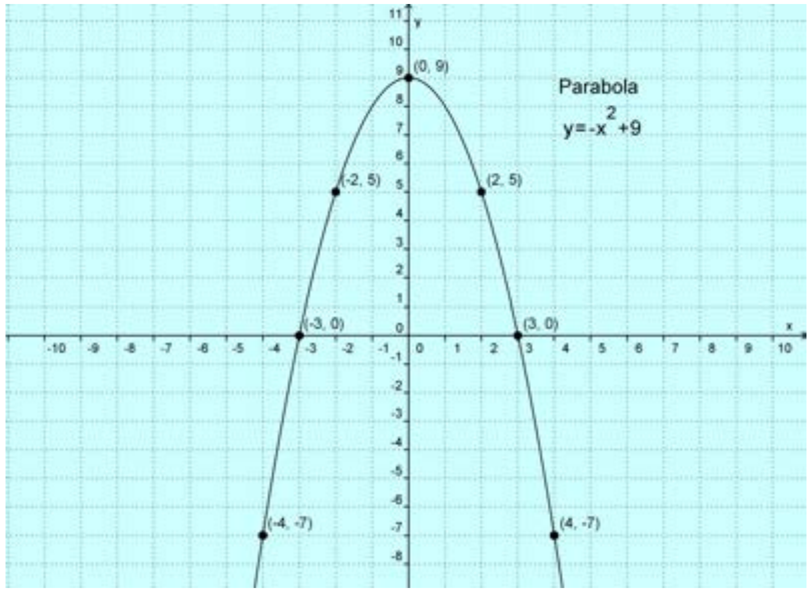
Figure 2.3 Circle

This is a circle that's graphed from the equation $x^2 + y^2 = 100$

Let's try another example. This time let's use the equation $y = -(x^2) + 9$. Again, let's plug in some numbers and begin plotting points. Let's try x = 0, we then get y=9, so let's plot (0, 9). Now x = 1, and you get 8. The same goes with x = -1, so plot the two points (1, 8) and (-1, 8). Next try $x = \pm 2$, and you get the two ordered pairs (2, 5) and (-2, 5). A few more calculations gets us a few more ordered pairs to plot (3, 0), (-3, 0), (4, -7), (-4, -7). Connect these points with the best curve you can, and you'll discover you've drawn a parabola (*Figure 2.4*).

As you can see from the two examples, some equations are closed, e.g. the circle, and some are open, e.g. the parabola.

Figure 2.4 Parabola Graph



This is a graph of a parabola given by the equation $y = -(x^2) + 9$.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/introduction-to-graphing/graphing-equations/

Distance Formula and Midpoints of Segments

The distance formula and the midpoint formula give us the tools to find important information about two points.

KEY POINTS

- The Pythagorean Theorem is a very powerful theorem relating the lengths of the three sides of a triangle. This theorem tells us that if c is the hypotenuse and a and b are the other two sides, $c^2 = a^2 + b^2$.
- Using the Pythagorean Theorem and two points (x1, y1) and (x2, y2), we can derive the distance formula: $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$.
- The midpoint of a line segment given by two points (x1, y1) and (x2, y2) is $(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2})$.

The Distance Formula

Let's say you have two points, for example (2, 4) and (5, 8) and want to know the **distance** between these two points. You could always take a ruler and attempt to find the distance based based on the units of the graph, but this is not accurate, nor will you always have access to a ruler. A much more accurate way to get the distance

between two points is to use geometry and the Pythagorean Theorem. This theorem states that in any right triangle, the square of the hypotenuse is equal to the sum of the squares of the other two sides (*Figure 2.5*).

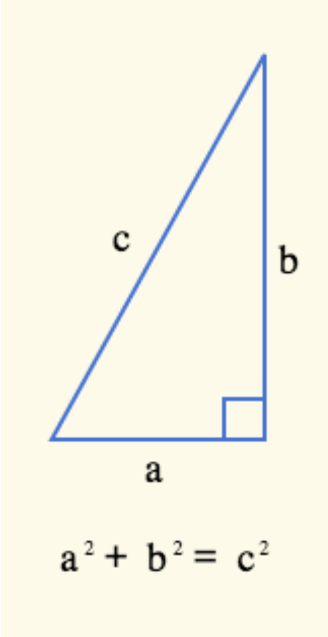


Figure 2.5
Pythagorean Theorem
The Pythagorean Theorem states that the square of the hypotenuse is equal to the sum of the squares of the other two sides.

Let's say that there are two dots on a coordinate plane. How would you find the distance between the two without a ruler? Hint: draw a right triangle. Suppose you have two points, (x_1, y_1) and (x_2, y_2), and suppose that the length of the straight line between them is d. You can derive the distance formula by noticing that you can follow the following path between any two points to obtain a right triangle: start at point one, change x (keep y constant) until you're directly above or below point two, and then alter y and keep x constant until you're at point two.

Using the **Pythagorean Theorem** and *Figure 2.6*, it becomes evident that

$d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 22$

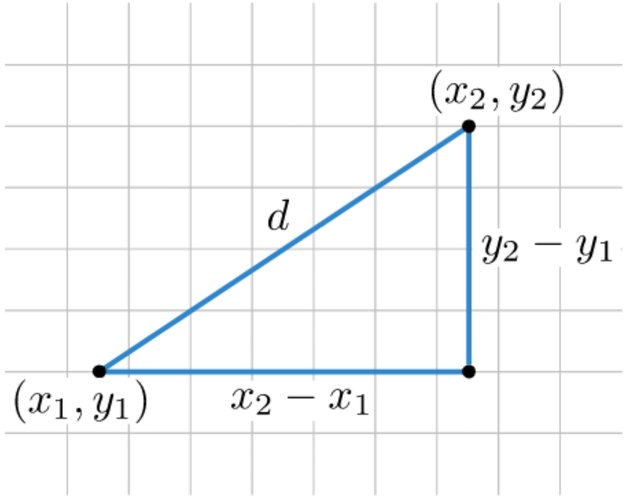


Figure 2.6
Distance Formula
The distance formula between two points, (x_1, y_1) and (x_2, y_2), shown as the hypotenuse of a right triangle

and solving for d by taking the square root of both sides, we have the full distance formula:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Using this formula and our original two points (2, 4) and (5, 8), we can plug those values into $x_1,\ x_2,\ y_1\ and\ y_2$ to get the distance.

$$d = \sqrt{(5 - 2)^2 + (8 - 4)^2} = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$$

Midpoint of a Line Segment

In geometry, the **midpoint** is the middle point of a line segment. It is equidistant from both endpoints. If you have two points, (x_1, y_1) and (x_2, y_2), the midpoint of the segment connecting the two points can be found with the formula $(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2})$

By looking at each coordinate, you can see that the x coordinate is halfway between x_1 and x_2, and the y coordinate is halfway between y_1 and y_2 (*Figure 2.7*).

Using our example points from above, (2, 4) and (5, 8), and the midpoint formula, we see that the midpoint of the line connecting these two points is $(\frac{2+5}{2}, \frac{4+8}{2}) = (\frac{7}{2}, 6)$

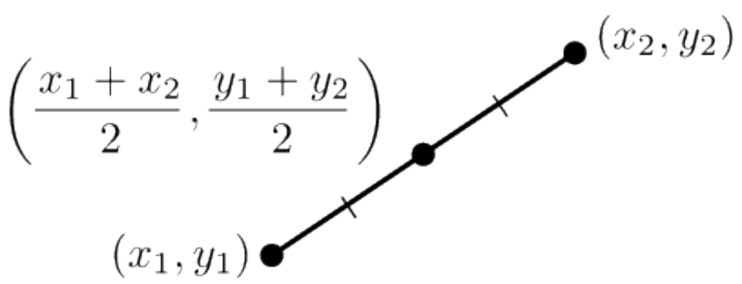


Figure 2.7
Midpoint of a Line Segment
The equation for a midpoint of a line segment with end points (x_1, y_1) and (x_2, y_2)

Source: https://www.boundless.com/algebra/graphs-functions-and-models/introduction-to-graphing/distance-formula-and-midpoints-of-segments/

Circles

The equation for a circle is just an extension of the distance formula.

KEY POINTS

- A circle is defined as the set of points that are a fixed distance from a center point.
- The distance formula can be extended directly to the definition of a circle by noting the radius is simply the distance between the center of a circle and the edge.
- The general equation for a circle, centered at (a, b), with radius r is the set of all points (x, y) such that $(x-a)^2 + (y-b)^2 = r^2$.

A **circle** is defined as the set of points that are a fixed distance from a center point. This definition is expressed in the way we draw circles. We pick a point as the center and then use some mechanical means to rotate a drawing utensil around that point. Of course our drawings are always approximations of the shape we think of as a circle. Drawings are only as accurate as the hardness and thickness of our tools will allow. For instance if you blow (*Figure 2.8*) up large enough on your computer screen you might see it is composed of adjacent colored squares and rectangles on your screen. If you used software that maintained the scale of the points as you magnified

the circle then eventually it would start to look like a line. This is the reason that we perceive roads as being straight even though we know the Earth is round.

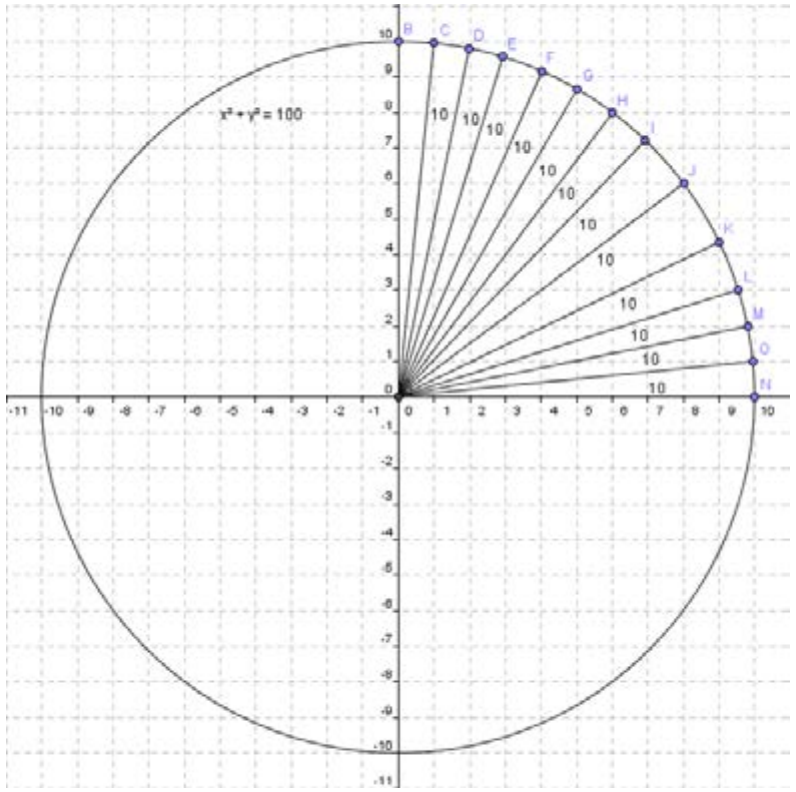


Figure 2.8 Circle

Graphed is a circle with radius 10 centered on the origin.

Since we know a circle is the set of points a fixed distance from a center point, let's look at how we can construct a circle in a Cartesian coordinate plane with variables x and y. To find a formula for this, suppose that the center is the point (a,b). According to the distance formula, the distance c from the point (a,b) to any other point (x, y) is $c^2 = (x - a)^2 + (y - b)^2$. The radius of the circle, r, is the distance between the center of the circle and any point along the edge. Therefore, the general equation for a circle, centered at (a, b), with **radius** r is the set of all points (x, y) such that $(x - a)^2 + (y - b)^2 = r^2$.

Functions: An Introduction

https://www.boundless.com/algebra/graphs-functions-and-models/functions-an-introduction/

Functions and Their Notation

A function maps a set of inputs onto permissible outputs, and each input maps onto one and only one output.

KEY POINTS

- Functions are relations between a set of inputs and a set of permissible outputs with the property that each input is related to exactly one output.
- Typically functions are named with a single letter, most commonly f, g and h. A function takes the form f(x) for one input variable, but it can take any number of variables, e.g. f(x, y, z).
- Functions can be thought of as a machine in a box open on two ends. You put something in one end, something happens to it in the middle, and something pops out the other end.

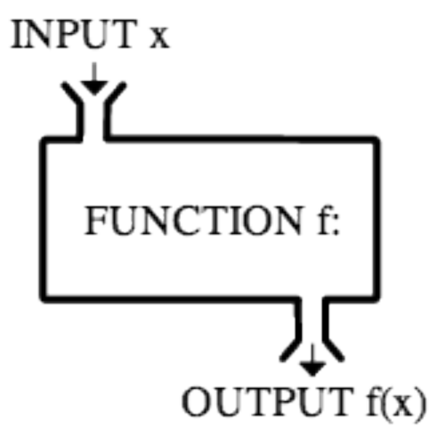
What Are Functions?

In mathematics, a **function** is a **relation** between a set of inputs and a set of permissible outputs with the property that each input is related to exactly one **output**. An example is the function that relates each real number x to its square x^2. Functions are typically named with a single letter, most typically f, so we'll call this function f. The output of a function f corresponding to an input x and is denoted by f(x) (read "f of x"). In this example, if the input is −3, then the output is 9, and we may write f(−3) = 9. The input variable(s) are sometimes referred to as the argument(s) of the function.

In the case of a function with just one input variable, the input and output of the function can be expressed as an ordered pair, ordered so that the first element is the input, the second the output. In the example above, $f(x) = x^2$, we have the ordered pair (−3, 9). If both the input and output are real numbers, this ordered pair can be viewed as the Cartesian coordinates of a point on the graph of the function.

Another commonly used notation for a function is f:X→Y, which reads as saying that f is a function that maps values from the set X onto values of the set Y.

Figure 2.9 Function Machine



A function f takes an input x and returns an output f(x). One metaphor describes the function as a "machine" or "black box" that for each input returns a corresponding output.

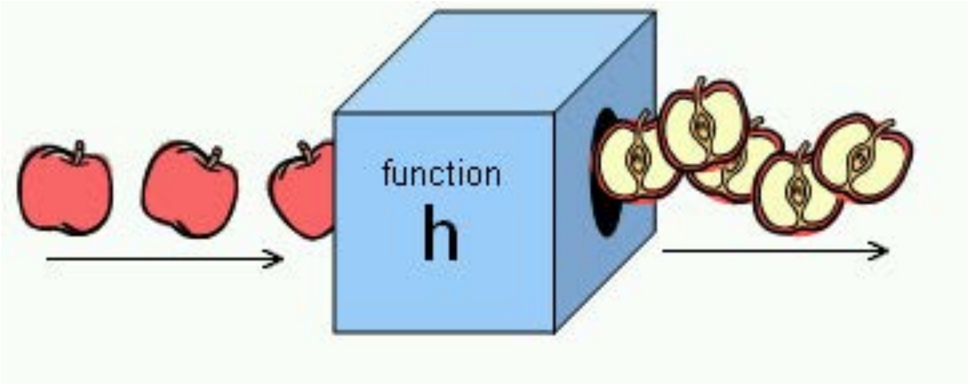
Functions As a Box

Functions are often described as a machine in a box open on two ends. You put something in one end, something happens to it in the middle, and something pops

out the other end. The function is the machine inside, and it's defined by what it does to whatever you give it (*Figure 2.9*).

Let's say the machine has a blade that slices whatever you put into it in two and sends one half out the other end. If you put in a banana, you'd get back half a banana. If you put in an apple, you'd get back half an apple (*Figure 2.10*).

Figure 2.10 Fruit Halving Function

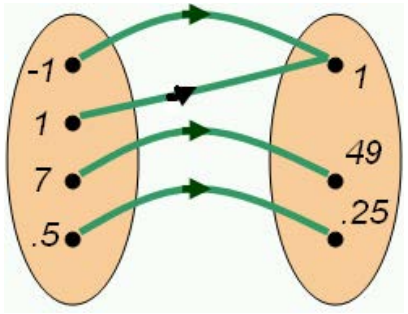


This shows a function that takes a fruit as input and releases half the fruit as output.

You may wonder what happened to the other half of the piece of fruit, but since this is algebra, the things that go in and come out of functions will be numbers, so the box simply fills up with numbers and will not break. Let's define the function to take what you give it and cut it in half, that is, divide it by two. If you put in 2, you'd get back 1. If you put in 57, you'd get back 28.5. The function machine allows us to alter expressions. In this example, $f(x) = \frac{1}{2}x$.

Functions As a Relation

Figure 2.11 Mapping of a Function



This shows some of the potential input to output of a function. For instance, -1 and 1 both map to the value 1, 7 to 49, and 0.5 to 0.25. The function is f(x)=x*x.

Functions can also be thought of as a subset of relations. A relation is a connection between numbers in one set and numbers in another (*Figure 2.11*). In other words, each number you put in is associated with each number you get out. The difference is that in a function, every input number is associated with exactly one output number, whereas in a relation, an input number may be associated with multiple or no output numbers. This is an important fact about functions that cannot be stressed enough: every possible input to the function must have one and only one output. All functions are relations, but not all relations are functions.

Graphing Functions

Graphs are a great visual representation of functions, showing the relation between the input and the output as a line or curve.

KEY POINTS

- Functions have an independent variable and a dependent variable. Typically x is the independent variable and y the dependent variable.
- As you choose any valid value for the independent variable, the dependent variable is determined by the function and you will always get the same result.
- To graph a function, choose some values for the independent variable, x, plug them into the function to get a set of ordered pairs (x, y), and plot these on the graph. Then connect the points to best match how the points are arranged on the graph. Make sure you have enough points.

Independent and Dependent Variables

Functions have an **independent variable** and a **dependent variable**. When we look at a function such as $f(x) = \frac{1}{2}x$, we call the variable that we are changing, in this case x, the independent variable. We assign the value of the function to a variable, in this case y, that we call the dependent variable. The reason that we say that x is independent is because we can pick any value for which the function is defined, in this case the real numbers R, as an input into the function. Once we pick the value of the independent variable the same result will always come out of the function. We say the result is assigned to the dependent variable, since it depends on what value we placed into the function.

It should be noted that you can switch the dependent and independent variable. Do this by equating y with our function $y = \frac{1}{2}x$, then $2y = 2(\frac{1}{2}x)$, then $2y = x$, and so $g(y) = 2y$. The independent variable is now y and the dependent variable x.

Graphing Functions

Let's start with a rather simple function, $f(x) = 5 - \frac{5}{2}x$.

How can we begin **graphing** this function? Start by choosing a few values for the independent variable, x. Say 0 and 2. Next, plug these values into the function, and we get the ordered pairs (0, 5) and (2, 0). Now this function is that of a line, so simply connect the two points, extend them in either direction past the points to infinity, and we have our graph (*Figure 2.12*).

Now let's choose a more interesting function, $f(x) = x^3 - 9x$.

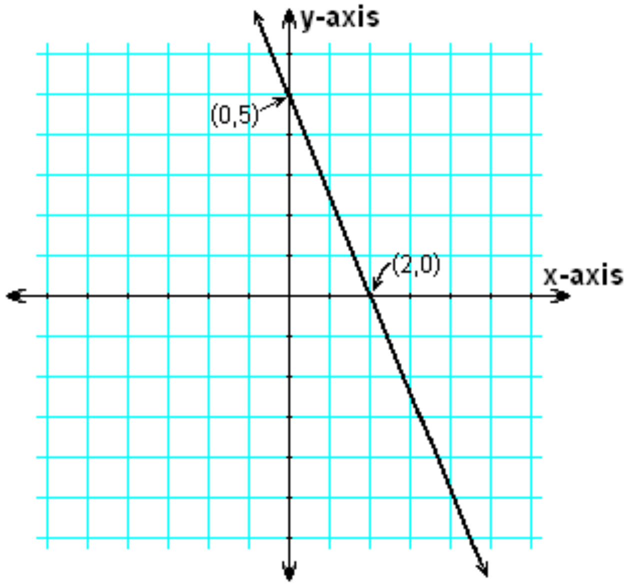


Figure 2.12 Graph of a Line

This is a graph of the line with function $f(x) = 5 - 2.5x$.

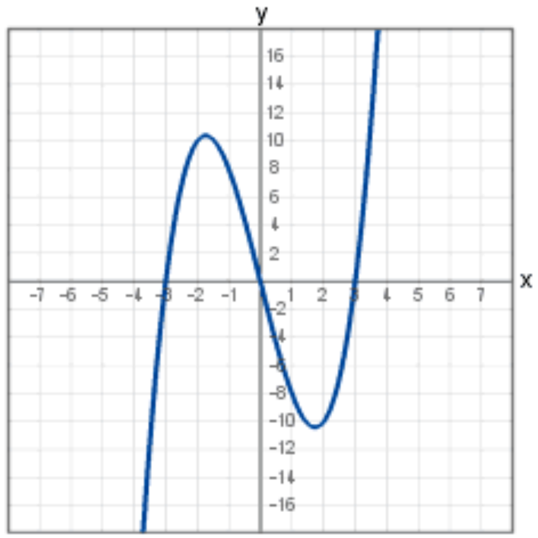


Figure 2.13 Graph of a Function

This is the graph of the function $f(x) = x^3 - 9x$.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/functions-an-introduction/graphing-functions/

Again, start by choosing a few values for the independent variable, x. This function is a little more complicated than a line, so we'll need to choose some more points. Say

$x = \{0, \pm 1, \pm 2, \pm 3, \pm 4\}$.

Next, plug these values into the function, $f(x) = x^3 - 9x$, to get a set of ordered pairs, in this case we get the set of ordered pairs:

$\{(-4, -28), (-3,0)(-2,10), (-1,8), (0,0), (1, -8), (2, -10), (3,0), (4,28)\}$.

Next place these points on the graph, and connect them as best as possible with a curve. If you don't have enough points to be sure about what the graph should look like, simply calculate some more! The graph for this function is (*Figure 2.13*).

Finding Domains of Functions

The domain of a function is the set of all possible input values that produce some output value.

KEY POINTS

- Given a function f:X→Y, the set X is the domain of f, and the set Y is the codomain of f.
- The domain of a function f is all of the values for which the function is defined. For instance, $\sqrt{x}$ is not defined when x<0.
- To find the domain of a function f, you must find the values for which f is not defined.

What Is the Domain of a Function?

The **domain** of a function is the set of "input" or argument values for which the function is defined. The domain is shown in the left oval in (*Figure 2.14*). The function provides an "output" or value for each member of the domain. For instance, the domain of $f(x) = x^2$ is the set of all real numbers, R, as every real number you put into f will give an output, namely x^2.

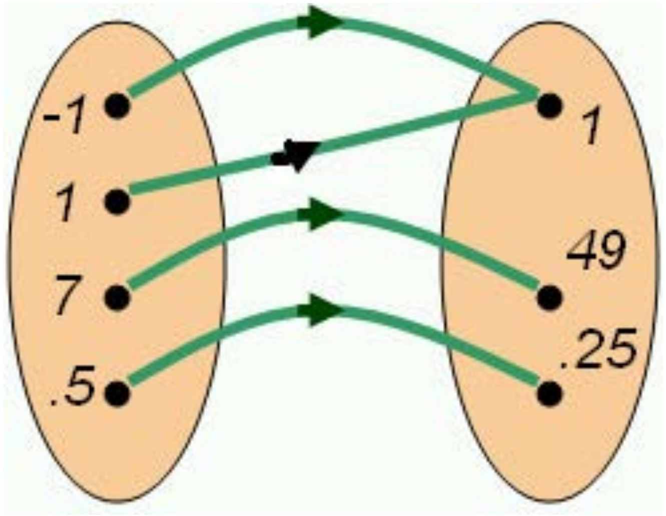


Figure 2.14 Mapping of a Function

The oval on the left is the domain of the function f, and the oval on the right is the range.

Given a function f:X→Y, the set X is the domain of f; the set Y is the codomain of f. In the expression f(x), x is the argument, and f(x) is the value. One can think of an argument as an input to the function and the value as the output.

It is important to note that not all functions have the set of real numbers as their domain. For instance, the function $f(x) = \dfrac{1}{x}$ is not defined for x=0, because you cannot divide a number by 0. In this case, the domain of f is the set of all real numbers except 0. That is, $x \neq 0$. So the domain of this function is $\mathrm{R} - \{0\}$.

What about the function $f(x) = \sqrt{x}$? In this case, the square root of a negative number is not defined, and so the domain is the set of all real numbers with $x \geq 0$.

Finding the Domain of a Function

In order to find the domain of a function, if it isn't stated to begin with, we need to look at the function definition to determine what values are disallowed. For instance, we know that you cannot take the square root of a negative number, and you cannot divide by 0. With this knowledge in hand, let's find the domain of the function

$$f(x) = \frac{1}{\sqrt{x-1}-2} + x.$$

First, we know we cannot divide by 0, so any value of x that causes a division by 0 is disallowed in the domain. In this example, this occurs when $\sqrt{x-1} - 2 = 0$. Solving for x, this happens is when x=5, so we know that $x \neq 5$. We also know we can't take the square root of a negative number. This means that $x - 1 > 0$ or $x > 1$. So this function's domain is the set of all real numbers such that x > 1 and $x \neq 5$. Therefore, to find what values are not in the domain, you must find the values where the function (or parts of it) is not defined.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/functions-an-introduction/finding-domains-of-functions/

Visualizing Domain and Range

All values in the domain are mapped onto values in the range that consist of one of four types of relations.

KEY POINTS

- Values in the domain map onto values in the range.
- There are four types of relations between the domain and range, one-to-one, one-to-many, many-to-one, and many-to-many.
- The horizontal and vertical line test can help determine the type of relation between the domain and range.

Domain and Range

As stated in a previous section, the **domain** of a function is the set of 'input' numbers for which the function is defined. The domain is part of the definition of a function. In most algebra formulas, x is usually the variable associated with domain. For example, the domain of the function $f(x) = \sqrt{x}\ is\ x \geq 0$.

The **range** of a function is the set of results, or solutions, to the equation for a given input. A true function only has one result for every domain. In most algebra formulas, y is usually the variable

associated with Range. As such, it can also be expressed f(x), which says that its value is a function of x. For instance, the function $f(x) = x^2$ has a range of $f(x) \geq 0$, because the square of a number is always positive.

In taking both domain and range into account, a function is any mathematical formula that produces one and only one result for each input. Hence, it can be said that in a valid function, domain (x) and range (y) have a many to one correspondence so that every given domain value has one and only one range value as a result, but not necessarily vice versa. This makes sense since results can repeat, but inputs cannot.

Types of Relations

There are four types of relations, one-to-one, one-to-many, many-to-one, and many-to-many. These types describe how inputs and outputs are related.

1. One-to-one means that every input has one unique output, and every output has one unique input. This kind of relation is a function. An example of this relation is $f(x) = \pm 1$.

2. Many-to-one means that multiple inputs can map onto the same output, but each input still has only one output. This kind of relation is a function, as each input only has one output. An example of a many-to-one function is $f(x) = x^2$, as with both x=-1 and x=1, f(x) = 1. (*Figure 2.15*) is also an example of a many-to-one relation.

3. Many-to-many means that multiple inputs map onto the same output, and inputs have multiple outputs. This relation is not a function. An example of this relation is $f(x) = \pm x^2$, for instance, $x = 1$ and $x = -1$ has $f(x) = \pm 1$.*Figure 2.16* is also an example of a many-to-many relations.

The line test is a test to see the type of relation. The horizontal line test checks if multiple input values have the same output value. To perform a horizontal line test take a horizontal line and if it intersects with two or more points, then the test 'fails'. The vertical line test checks if a single input value has multiple output values. To perform a vertical line test take a vertical line and if it intersects with two or more points, then the test 'fails'. One-to-one has both horizontal and vertical line tests passing. One-to-many has the

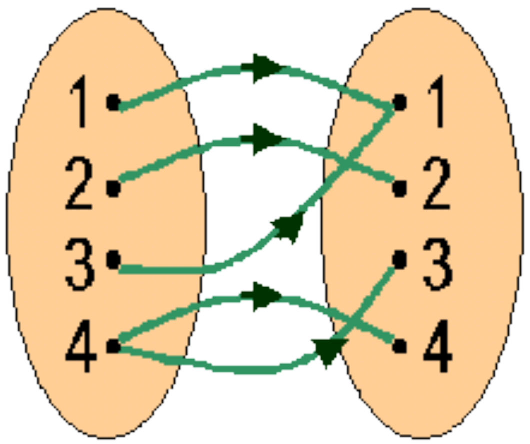


Figure 2.15 Many to Many Domain and Range

This is a many-to-many visualization of a domain and range. Because it is many-to-many it is not a valid function.

horizontal test passing and the vertical test failing. Many-to-one has the horizontal test failing and the vertical test passing. Many-to-many has both tests failing. An important result of these tests is that any graph where the vertical line test fails, then the graph is not a function.

Visualizing Domain and Range

The domain and range can be visualized by *Figure 2.15* and *Figure 2.16*. One can also use a graph, such as the graph for $f(x) = x^2$ in *Figure 2.17*. It is important to note that a graph is not always perfect, if for instance you are zoomed in on the graph where all values are non-negative, however for larger values of x, the graph turns around and becomes negative.

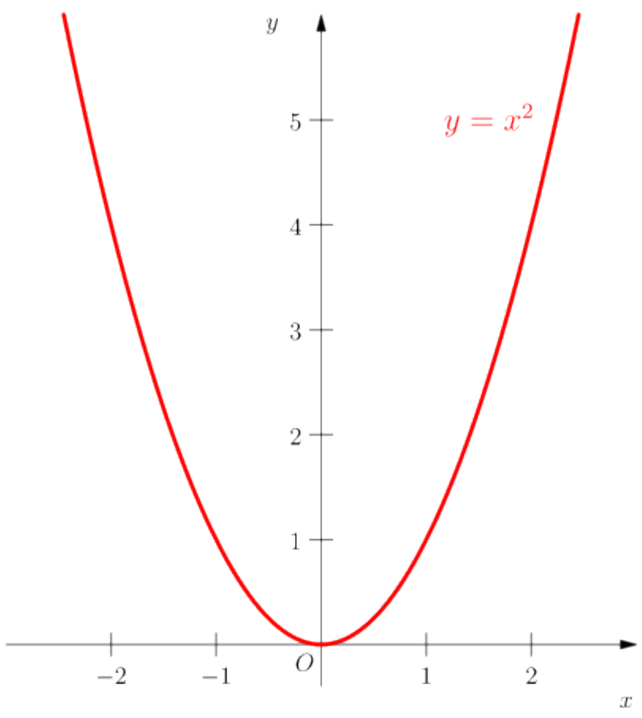


Figure 2.17
Domain and Range Visualization
Using graphs, you can visualize the domain and range, noting that the range is always a non-negative number.

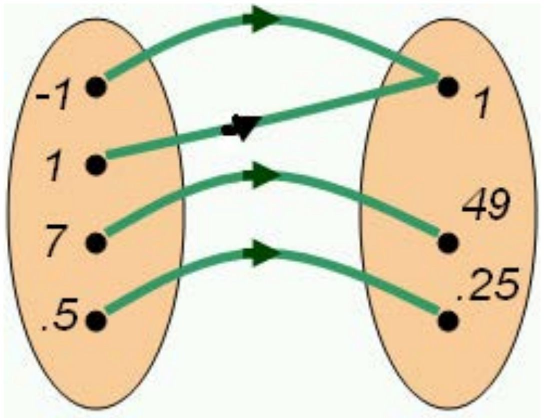


Figure 2.16
Mapping of a Function
This is a many-to-one visualization of a domain and range. Because it is many-to-one, this is a valid function.

The Linear Function f(x) = mx + b and Slope

The linear function f(x) = mx + b defines a line in a 2D coordinate system with slope m and y-intercept b.

KEY POINTS

- A linear function is an algebraic equation in which each term is either a constant or the product of a constant and (the first power of) a single variable. The typical linear function in slope-intercept form is y = mx+b where m and b are constants.
- The y-intercept, b, is the y-coordinate of the location where the line crosses the y-axis, which can be found by setting x to 0.
- The slope, m, is the change in the vertical distance of a line on a coordinate plane over the change in horizontal distance. The slope of the line is $m = \frac{y_2 - y_1}{x_2 - x_1}$.

What is a Linear Function?

A linear function is an algebraic equation in which each term is either a constant or the product of a constant and (the first power of) a single **variable**. Linear equations can have one or more variables. Linear equations occur with great regularity in applied mathematics. A common form of a linear function in the independent variable x and the dependent variable y is $y = mx + b$ where m and b are constants.

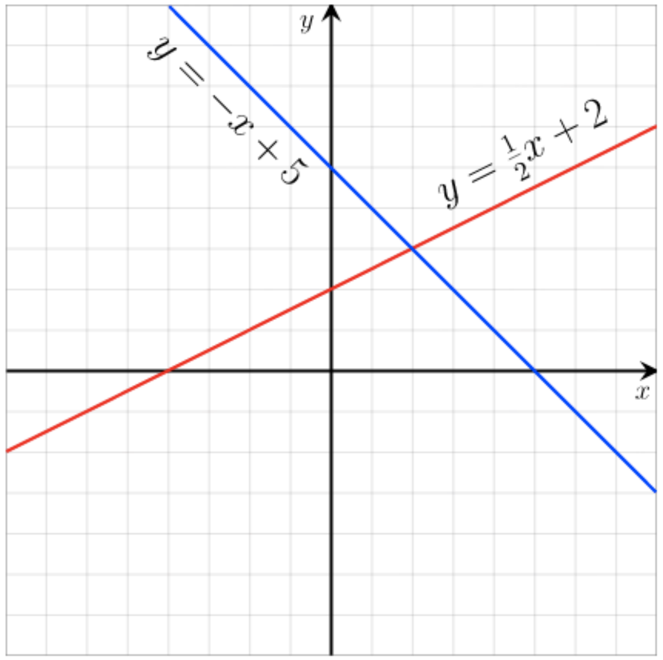
The origin of the name "linear" comes from the fact that the set of solutions of such an equation forms a straight line in the plane. In this particular equation, the constant m determines the slope or gradient of that line, and the constant term b determines the point at which the line crosses the y-axis, otherwise known as the y-intercept. (*Figure 2.18*) shows a graph of a couple of linear functions.

The y-intercept and the Slope

The y-intercept, which is b in the standard form, is the y-coordinate of the location where the line crosses the y-axis. This can be seen by letting x = 0, which immediately gives y = b. It may be helpful to think about this in terms of y = b + mx; where the line passes through the point (0, b) and extends to the left and right at a slope of m. Vertical lines having undefined slope, cannot be represented by this form. Vertical lines have the form x = c for some constant c.

It's also possible to find the x-intercept of the function, the x-coordinate where the line crosses the x-axis, but it isn't immediately clear from the intercept-slope form. To find the x-intercept, simple

Figure 2.18 Linear Function Graph



An example of two linear functions, y = -x + 5 and y = 0.5x+2. Both functions graph as straight lines, but their y-intercepts and slopes are determined through their functional terms. The positive (red) and negative (blue) slope terms change the orientation of the lines relative to each other.

solve for x to get $x = \frac{y-b}{m} = \frac{y}{m} - \frac{b}{m}$ and so you can see that by letting y = 0, the x-intercept is therefore $-\frac{b}{m}$.

The slope, m, is the change in the vertical distance of a line on a coordinate plane over the change in horizontal distance. In other words, it is the "rise" over the "run" or the steepness of a line. Lines that have the same slope are parallel lines, and these lines will never touch or cross. Slope is computed by measuring the change in vertical distance divided by the change in horizontal difference, in order to compute this you'll need two points on the line, say (x_1, y_1) and (x_2, y_2), such that $x_2 > x_1$. From these two points, and the knowledge that the slope is the change in the vertical distance over the change in the horizontal distance, it's easy to see the slope $m = \frac{y_2 - y_1}{x_2 - x_1}$.

Finding the Intercept and Slope

The linear function won't always be in slope-intercept form, or sometimes you'll need to find the function from a pair of points. If you have a function that is not in the slope-intercept form, to get it into that form, simply solve for y. For example, if you have the function $\frac{3y - 15}{6} = 2x$ solve for y by first multiplying both sides by 6 to get $3y - 15 = 12x$, add 15 to both sides for $3y = 12x + 15$, and finally divide both sides by 3 to put it into slope-intercept form $y = 4x + 5$, which tells us the y-intercept is 5, and the slope is 4.

Another possibility is that you are simply given two points, say (1, 3) and (4, 0). First find the slope from the slope equation $m = \frac{0-3}{4-1} = \frac{-3}{3} = -1$. Now that we know the slope, choose one of the two points and plug it into the slope-intercept form. Let's choose (1, 3), plug this in to get $3 = -1 \times 1 + b$, and solve for b to get

b = 4. Therefore the slope-intercept form for the general function is $y = -1x + 4$.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/functions-an-introduction/the-linear-function-f-x-mx-b-and-slope/

Applications of Linear Functions and Slope

Linear functions apply to real world problems that involve a constant rate.

KEY POINTS

- If you know a real-world problem is roughly linear, such as the distance you travel when you go for a jog, all you need is two points and you can graph the function and make some assumptions as to what happens beyond the two points, so long as you maintain the same rate.
- The slope of a function is the same as the rate of change for the dependent variable. For instance, if you're graphing distance vs. time, then the slope is how fast your distance is changing with time, or in other words, your speed.
- When checking where two linear functions intersect, set them equal to each other and solve for the dependent variable, x.

Let's say that one day you decide to start training for a marathon. You start at your house, stretch, and look at your watch. It reads 6:00 pm. You plug in your headphones and begin to run around town. After a while, you realize you can't run anymore, and look at your watch. It reads 7:30 pm, and you're 7.5 miles from home. How fast was your average speed over the course of the run?

Figure 2.19 Trains Used in Applications

Trains are just one example of things that can be used with linear functions. For instance, to see when two trains travelling at constant rates towards each other meet is a simple linear function.

Our two variables are time and distance, and you have the data for two separate points. The first point is at your house, where your watch read 6:00, let's call this the beginning time and set it to 0. So our first point is (0, 0) as we hadn't run any distance yet. Let's think about our time in hours. So our second point is 1.5 hours later, and we ran 7.5 miles. So our second point is (1.5, 7.5). Our speed is simply the **slope** of the line connecting the two points. The slope, given by

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

becomes

$$m = \frac{7.5}{1.5} = 5 \text{ miles per hour.}$$

To graph this line, we need the y-intercept and the slope. We have just calculated the slope, but what's the y-intercept? Since we started at (0, 0), we can see that the y-intercept is 0. So our final function is $y = 5x$.

With this new function, we can now answer some more questions.

How many miles had we run after the first half hour?

If we kept running at the same pace for a total of 3 hours, how many miles will we have run? $26 = 5x$, so $x = \dfrac{26}{5} = 5.2$ hours.

There are many such applications for **linear equation**s. Anything that involves a constant rate of change can be nicely represented with a line with the slope. Indeed, so long as you have just two points, if you know the function is linear, you can graph it and begin asking questions! Just make sure what you're asking and graphing makes sense. For instance, in the marathon example, the domain is really only $x \geq 0$, since it doesn't make sense to go into negative time and lose miles!

Let's look at another application. Two trains start 200 miles apart. They are travelling towards each other. The first train is travelling 40 miles per hour, while the second train is travelling 60 miles per hour. When and where do they meet? For simplicity we should put one of the trains starting at (0, 0), and so the other train must start at either (0, 200) or (0, -200) as it is 200 miles away. We'll choose (0, 200). The first train is travelling 40 miles per hour towards the other train, which is in the positive y direction, so the slope is positive 40. The second train is travelling towards the first at 60 miles per hour, which is in the negative y direction, so the slope is negative 60. Now we have the equations for both of the functions. Let $f(x) = 40x$ represent the first train and $g(x) = -60x + 200$ represent the second train. To find when they meet, simply set them equal to each other and solve for x.

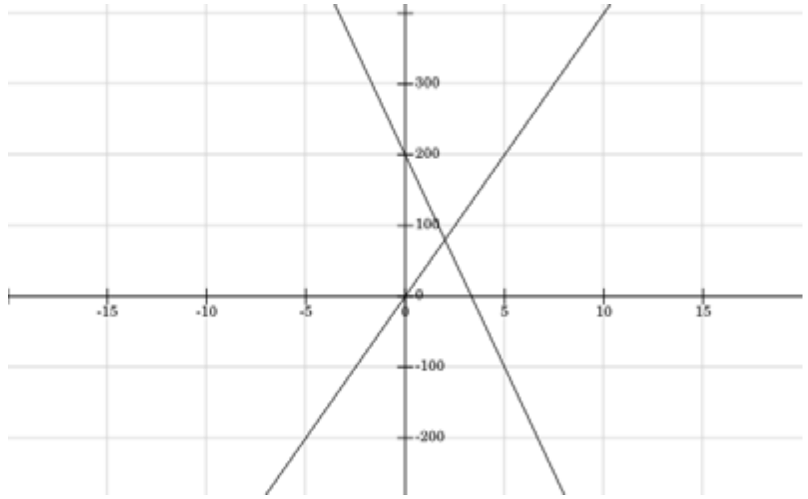
$$40x = -60x + 200$$

$$100x = 200$$

$$x = 2$$

They meet after 2 hours of travel. To find out where they meet, plug x=2 into one of the equations, say f(x). Thus $f(2) = 40 \times 2 = 80$ so they meet 80 miles away from where the first train started, or 120 miles from where the second train started. This can also be viewed by graphing the two equations and seeing their intersection ([Figure 2.20]()).

Figure 2.20 Two Trains Graphically



The intersection of f(x) = 40x and g(x) = -60x + 200 can be viewed graphically. This can be used to determine where they will cross.

Section 3

Modeling Equations of Lines

https://www.boundless.com/algebra/graphs-functions-and-models/modeling-equations-of-lines/

Slope-Intercept Equations

The slope-intercept form of a line, gives you the information to construct a quick and easy line using the slope, m, and the y-intercept, b.

KEY POINTS

- The slope-intercept form of a line is given by $y = mx + b$ where x and y are variables, and m and b are constants. The constant m is the slope, and b is the y-intercept.
- The constant b is known as the y-intercept, when x=0, y = b, and the point (0, b) is the unique member of the line where the y-axis is 'intercepted' or crossed.
- The x-intercept, when y = 0, x=-b/m, assuming m is non-zero, and the point (-b/m, 0) is the unique member of the line where the x-axis is 'intercepted' or crossed.
- The constant m is known as the slope. Slope is the measure of how much a line moves up or down related to how much it moves left to right.
- If you have two points (x1, y1) and (x2, y2), where $\frac{y_2 - y_1}{x_2 - x_1}$.

One of the most common representations for a line on a graph is with the **slope**-intercept form. Such an equation is given by $y = mx + b$, where x and y are variables and m and b are constants. The constant m is known as the slope and b is known as the **y-intercept**. Note that if m is 0, then y = b, and thus y is a horizontal line. Note that this equation does not allow for vertical lines, since that would require that m be infinite, though a vertical line can be define by x = c for some constant c.

Intercepts

The constant b is known as the y-intercept. The equation $y = f(x) = mx + b$ has an infinite number of solutions. Points will be mapped with independent variable x on the horizontal axis, and y on the vertical. By assigning x to a value and evaluating, you get a single point (x, y). When x=0, y = b, and the point (0, b) is the unique member of the line where the y-axis is 'intercepted' or crossed.

There is also an x-intercept, but it is not immediately clear from the equation. To find the x-intercept, set y = 0 and solve for x to get $x = -\frac{b}{m}$, giving the point $(-\frac{b}{m}, 0)$ and this is the unique point where the x-axis is 'intercepted'.

Exceptions to these are the horizontal and vertical lines. Horizontal lines given by y = b can intercept the x-axis 0 or infinite times. It intersects an infinite number of times if b = 0, and 0 times otherwise. Similarly the vertical line given by x = c can intercept the

y-axis 0 or infinite times. It intersects an infinite number of times if c = 0, and 0 times otherwise.

Slope

The constant m is known as the slope. Slope is the measure of how much a line moves up or down related to how much it moves left to right. Slope is computed by measuring the change in vertical distance divided by the change in horizontal difference, i.e. $m = \frac{\Delta y}{\Delta x} = \frac{rise}{run}$. So if there are two points (x_1, y_1) and (x_2, y_2) where $x_1 \neq x_2$, then m = $\frac{y_2 - y_1}{x_2 - x_1}$

If a line goes up from left to right, then the slope has to be positive. For example, a slope of ¾ would have a "rise" of 3, or go up 3; and a "run" of 4, or go right 4. Both numbers in the slope are either negative or positive in order to have a positive slope.

If a line goes down from left to right, then the slope has to be negative. For example, a slope of -3/4 would have a "rise" of -3, or go down 3; and a "run" of 4, or go right 4. Only one number in the slope can be negative for a line to have a negative slope.

There are two special circumstances, no slope and slope of zero. A horizontal line has a slope of 0 and a vertical line has an undefined slope. Horizontal lines have the form y = b, where b is a constant. Vertical lines have the form x = c, where c is a constant.

Figure 2.21 shows numerous different lines with the same y-intercept at y=1, and a multitude of differing slopes.

Two separate line equations with the same slope are parallel lines. For example y = 2x+1 and y = 2x - 1 are parallel lines.

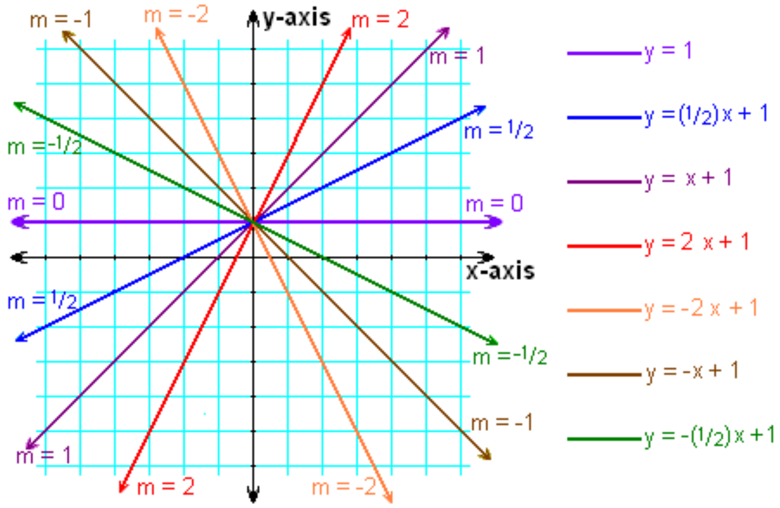


Figure 2.21
Differing Slopes
This is a graph of multiple lines with the same y-intercept, 1, and many differing slopes.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/modeling-equations-of-lines/slope-intercept-equations/
CC-BY-SA

Boundless is an openly licensed educational resource

Point-Slope Equations

The point-slope equation is another way to represent a line; to use the point-slope equation, only the slope and a single point are needed.

KEY POINTS

- The point-slope equation is given by (x_1, y_1) is any point on the line, and m is the slope of the line.
- The point-slope equation requires that there are at least one point and the slope. If there are two points and no slope, the slope can be calculated from the two points. Then use one of the two points in the final equation.
- The point-slope and slope-intercept equations are equivalent, and with a little bit of algebraic manipulation it can be shown that given a point $y_1 - mx_1$.

The **point-slope equation** is another way of describing the equation of a line. The point-slope form is great if you have the slope and only one point, or if you have two points and do not know what the y-intercept is. Given a slope, m, and a point (x_1, y_1), the point-slope equation is given by

$y - y_1 = m(x - x_1)$.

Changing between point-slope and slope-intercept forms is a simple task. If an equation is in a slope-intercept form, such as $y = mx + b$, to switch to point-slope form only one point is needed (x_1, y_1) that satisfies the slope-intercept form. By choosing either an x value or a y value, solving for the other variable, and plugging that information into the point-slope equation, using the same m, will get the other equation.

To show that these two equations are equivalent, choose a generic point (x_1, y_1). Plug in the generic point into the equation. The equation is now, $y_1 = mx_1 + b$, giving us the ordered pair, $(x_1, mx_1 + b)$. Then plug this point into the point-slope equation to get:

- $y - (mx_1 + b) = m(x - x_1)$
- $y - mx_1 - b = mx - mx_1$
- $y - mx_1 + mx_1 - b = mx - mx_1 + mx_1$
- $y - b = mx$
- $y - b + b = mx + b$
- $y = mx + b$

Therefore, the two equations are equivalent so long as any point on the line is chosen.

To reverse this process, from point-slope to slope-intercept follows a similar process. To switch from point-slope to slope-intercept, solve for y. For instance, in the generic case $y - y_1 = m(x - x_1)$, solve for y and multiplying m throughout. The equation ends up as $y = mx - mx_1 + y_1$.

If one rearranges equation and puts in parentheses, the result is

$y = mx + (y_1 - mx_1)$.

This equation looks very similar to slope-intercept, and indeed $y_1 - mx_1$ is a constant, and is the y-intercept of the equation.

It is interesting to note that, if the point chosen for the point-slope form has x_1 =0, then the equation comes out to be $y - y_1 = mx$. If we add y_1 to both sides, slope-intercept form is obtained. Thus, the slope-intercept form is just a basic case of the point-slope form with x_1 = 0.

Examples

If there is a single point (2, 1) and the slope is known to be -4. The point-slope equation for this line is $y - 1 = -4(x - 2)$. To switch this equation into slope-intercept form:

- $y - 1 = -4(x - 2)$
- $y - 1 = -4x + 8$
- $y = -4x + 9$

If there are two points, (2, 6) and (7, -1), first find the slope by remembering that given two points the equation of the slope is $m = \frac{y_2 - y_1}{x_2 - x_1}$. Therefore, the slope of this line is $m = \frac{-1-6}{7-2} = \frac{-7}{5}$.

Now choose any of the two points, such as (2, 6). Plug those points into the point-slope equation to get $y - 6 = \frac{-7}{5}(x - 2)$.

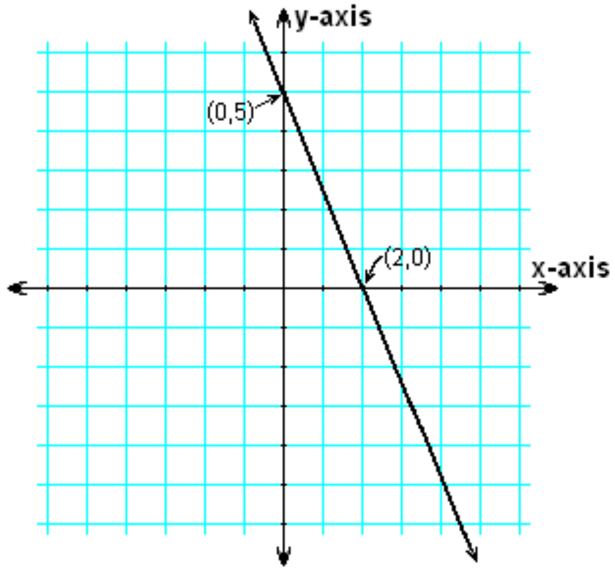


Figure 2.22 Graph of a Line

This is the graph of the line y-5=-2.5(x-0)

Parallel and Perpendicular Lines

Parallel lines never intersect; perpendicular lines intersect at right angles.

KEY POINTS

- Parallel lines have the same slope. A line is technically parallel with itself.
- Perpendicular lines slopes are negative reciprocals of each other. This means that one has m and the other -1/m.
- The distance between two parallel lines $d = \frac{|b_2 - b_1|}{\sqrt{m^2+1}}$.
- If the line f(x) is perpendicular to g(x), then f(x) is perpendicular to all lines that are parallel with g(x), and all lines parallel with f(x) are perpendicular to g(x).
- If line f(x) is perpendicular to g(x), and g(x) is perpendicular to h(x), then h(x) and f(x) are parallel.

Two lines in a plane that do not intersect or touch at a point are called **parallel lines**. The parallel symbol is ||. For example, given two lines

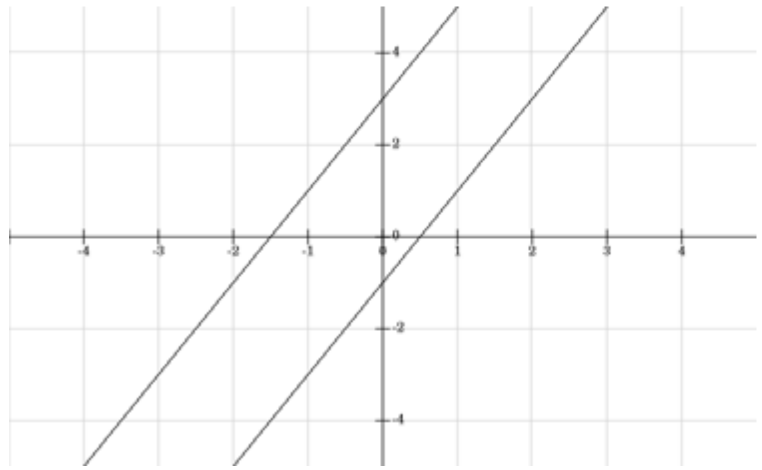


Figure 2.23 Parallel Lines

These are two parallel lines: f(x)=2x+3 and g(x)=2x-1.

$f(x) = m_1x + b_1$ and $g(x) = m_2x + b_2$, $f(x) \parallel g(x)$ states that the two lines are parallel to each other. Given two parallel lines f(x) and g(x), the following is true:

1. Every point on f(x) is located at exactly the same minimum distance from g(x).
2. Line f(x) is on the same plane as g(x) but does not intersect g(x), even assuming that the two lines extend to infinite in either direction.

In 2D, two lines are parallel if they have the slope. Recall that the slope intercept form $y = mx + b$ and the point slope form $y - y_1 = m(x - x_1)$ both contain information about the slope, namely the constant m. If two lines, say $f(x) = mx + b$ and $g(x) = nx + c$, are parallel, then n must equal m.

For example, in *Figure 2.23*, the two lines $f(x) = 2x + 3$ and $g(x) = 2x - 1$ are parallel since they have the same slope, m = 2.

Perpendicular Lines

Two lines are **perpendicular** to each other if they form congruent adjacent angles -- in other words, they are perpendicular if the angles at their intersection are right angles of 90 degrees (*Figure 2.24*). The perpendicular symbol is ⊥. For example given two lines, $f(x) = m_1x + b_1$ and $g(x) = m_2x + b_2, f(x) \perp g(x)$ states that the two lines are perpendicular to each other.

For two lines in a 2D plane to be perpendicular, their slopes must be negative **reciprocals** of one another. This means that if the slope of one line is m, then the slope of its perpendicular line is $\frac{-1}{m}$. In other words, the two slopes multiplied together must equal -1.

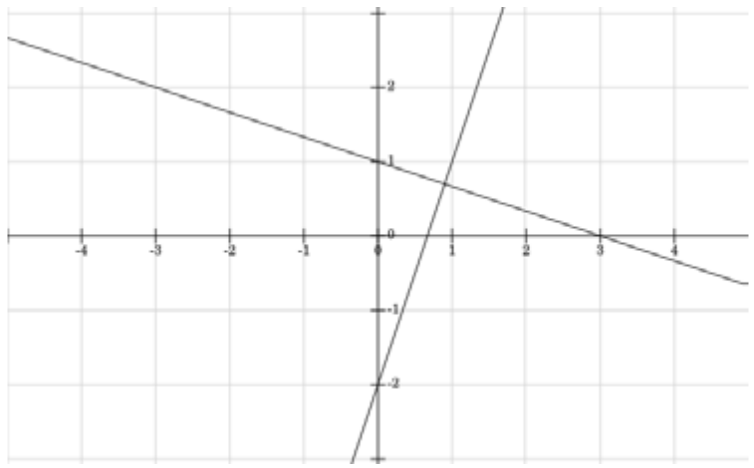


Figure 2.24 Perpendicular Lines

These two lines are perpendicular, where one line has f(x)=3x-2 and the other is g(x)=(-1/3)x +1. Note their slopes are negative reciprocals of each other.

For example, in *Figure 2.24*, the two lines are $f(x) = 3x - 2$ and $g(x) = \frac{-1}{3}x + 1$. Note that perpendicular lines have slopes that are negative reciprocals of each other. Additionally, if line f(x) is perpendicular to g(x), then f(x) is perpendicular to all lines that are parallel to g(x), and all lines parallel with f(x) are perpendicular to g(x).

If line f(x) is perpendicular to g(x), and g(x) is perpendicular to h(x), then h(x) and f(x) are parallel. However, this fact is only true in two dimensions!

Distance Between Two Parallel Lines

Let's say we want to know the distance between two parallel lines $y = mx + b_1$ and $y = mx + b_2$. The shortest distance between these two lines is the perpendicular line between them, thus any line with slope $\frac{-1}{m}$. We'll choose the perpendicular line to be $y = \frac{-x}{m}$. To find where the perpendicular line and the first parallel line intersect, set them equal to each other and solve for x to get:

- $mx_1 + b_1 = \frac{-x_1}{m}$
- $m^2x_1 + mb_1 = -x_1$
- $m^2x_1 + x_1 = -mb_1$

- $x_1(m^2+1) = -mb_1$
- $x_1 = \frac{-mb_1}{m^2+1}$

Now we have our x value, we need to plug this into the function to get y.

$$y_1 = \frac{-x_1}{m} = \frac{-1}{m}\frac{-mb_1}{m^2+1} = \frac{b_1}{m^2+1}$$

A similar argument can be used to find the point where the perpendicular line intersects with the second parallel line. This gives us the two points $(x_1, y_1) = (\frac{-mb_1}{m^2+1}, \frac{b_1}{m^2+1})$ and $(x_2, y_2) = (\frac{-mb_2}{m^2+1}, \frac{b_2}{m^2+1})$. To find the distance between these two points we use the distance formula:

- $d^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$
- $d^2 = (\frac{-mb_2 + mb_1}{m^2+1})^2 + (\frac{b_2 - b_1}{m^2+1})^2$
- $d^2 = \frac{m^2b_2^2 - 2m^2b_1b_2 + m^2b_1^2 + b_2^2 - 2b_1b_2 + b_1^2}{(m^2+1)^2}$
- $d^2 = \frac{(m^2+1)b_2^2 - 2(m^2+1)b_1b_2 + (m^2+1)b_1^2}{(m^2+1)^2}$
- $d^2 = \frac{b_2^2 - 2b_1b_2 + b_1^2}{m^2+1}$
- $d = \sqrt{\frac{b_2^2 - 2b_1b_2 + b_1^2}{m^2+1}}$
- $d = \frac{|b_2 - b_1|}{\sqrt{m^2+1}}$

Linear Mathematical Models

Linear mathematical models take real world applications and describe them with lines.

KEY POINTS

- A mathematical model describes a system using mathematical concepts and language.
- Linear mathematical models can be described with lines. For instance, a car going 50 mph, has traveled a distance represented by y = 50x, where x is time in hours and y is miles.
- Real world applications can be modeled with multiple lines. If two trains travel toward each other—one at 50 mph and the other at 80 mph—and they initially start at 350 miles away, the two lines are y=50x and y=350-80x. The point where the two lines intersect is the point where the trains meet.

A **mathematical model** is a description of a system using mathematical concepts and language. Mathematical models are used not only in the natural sciences (such as physics, biology, earth science, and meteorology) and engineering disciplines (such as computer science and artificial intelligence), but also in the social sciences (such as economics, psychology, sociology, and political science). Physicists, engineers, statisticians, operations research analysts, and economists use mathematical models most extensively. A model may help explain a system, study the effects of different components, and make predictions about behavior.

Dead Reckoning

Many everyday activities require the use of mathematical models, perhaps unconsciously. An example is seen in predicting the position of a vehicle based on its initial position, direction, and speed of travel. This is known as "dead reckoning" when used more formally (*Figure 2.25*). This type of mathematical modeling does not necessarily require formal mathematics; animals have been shown to use dead reckoning.

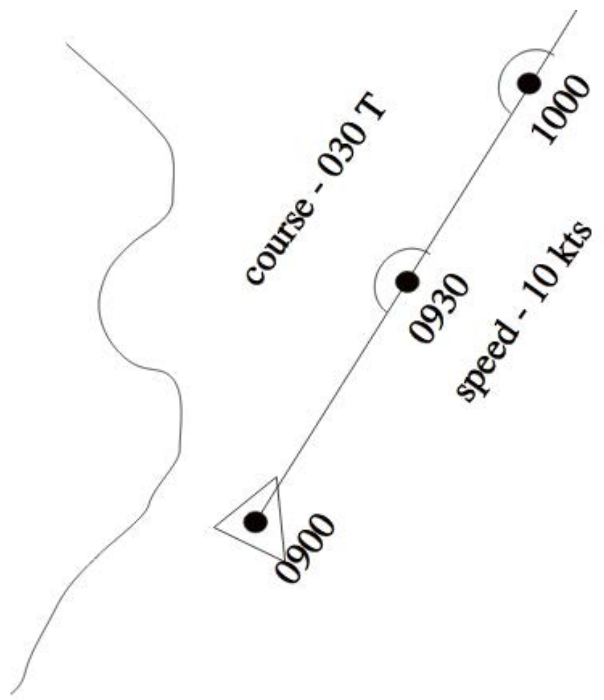


Figure 2.25 Dead Reckoning

This is an example of dead reckoning, which uses a linear model to predict how far one has traveled over time.

Example 1

One difficulty with mathematical models lies in translating the real world application into an accurate mathematical representation. Linear models can be used for problems concerned with straight lines, such as how far a vehicle has traveled from its initial position, moving at 50 mph. This can be easily modeled with a line y = 50x, where x is in hours. For instance, if you travel for 3.5 hours, then:

$y = 50(3.5) = 175 \ miles$

Example 2

It's also possible to model multiple lines at once. For instance, you have two trains, A and B. At the beginning, the A and B are 325 miles away from each other. Train A is traveling towards B at 50 miles per hour; B is traveling towards A at 80 miles per hour. First let's model the position of the two trains, using the initial location of A as the reference point. The origin (0, 0) is where A starts. B is 325 miles away from A at the beginning, so its position is (0, 325). Since A is traveling towards B, which has a greater y value, A's slope must be positive, equal to its speed, 50. B is traveling towards A, which has a lesser y value, giving B a negative slope: -80.

The two lines are thus $y_A = 50x \ and \ y_B = -80x + 325$.

This model allows for several observations—the two trains will meet where the two lines intersect.

Setting $y_A = y_B$ we have $50x = -80x + 325$.

Solving for x gives us x =2.5.

The two trains meet after 2.5 hours. To find where this is, plug 2.5 into either equation.

Plugging it into the first equation gives us $50(2.5) = 125$, which means it meets after A travels 125 miles.

For multiple observations that are roughly linear, you can use linear regression to create a line that best fits the data. This line allows linear modeling to provide new observations about the data.

Fitting a Curve

Curve fitting with a line attempts to draw a line so that it "best fits" all of the data.

KEY POINTS

- Curve fitting is useful for finding a curve that best fits the data, so you can see how the data is roughly spread out as well as make observations about data you don't have.
- Linear regression attempts to graph a line that best fits the data available.
- Ordinary least squares approximation is a type of linear regression that minimizes the sum of the squares of the difference between the approximated value (from the line), and the actual value.
- The slope of the line that approximates n data points is given by $m = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j}{\sum_{i=1}^{n} (x_i^2) - \frac{1}{n} (\sum_{i=1}^{n} x_i)^2}$.
- The y-intercept of the line that approximates n data points is given by $b = \frac{1}{n} \sum_{i=1}^{n} y_i - m \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{y} - m\bar{x}$.

Curve fitting is the process of constructing a curve, or mathematical function, that has the best fit to a series of data points, possibly subject to constraints. Curve fitting can involve either interpolation, where an exact fit to the data is required, or smoothing, in which a "smooth" function is constructed that approximately fits the data. Fitted curves can be used as an aid for data visualization, to infer values of a function where no data are available, and to summarize the relationships among two or more variables. Extrapolation refers to the use of a fitted curve beyond the range of the observed data, and is subject to a greater degree of uncertainty since it may reflect the method used to construct the curve as much as it reflects the observed data.

In this section, we will only be fitting lines to data points, but it should be noted that one can fit polynomial functions, circles, piecewise functions, and any number of functions to data and it is a heavily used topic in statistics.

Linear Regression

Linear regression is an approach to modeling the linear relationship between a dependent variable y and an independent variable x. Under linear regression, a line of the form $y = mx + b$ is found that "best fits" the data.

The simplest and perhaps most common linear regression model is the ordinary **least squares approximation**. This approximation attempts to minimize the sums of the squared distance between the line and every point. Suppose there are n points, our equation for

the line is $y = mx + b$. Given a point (x_i, y_i), using the same independent variable and plugging it into the equation we get another point $(x_i, \hat{y}_i)$, where $\hat{y}_i = mx_i + b$ and $\hat{y}_i$ and y_i may or may not be equal, we don't know. What we're trying to minimize is the sum of the squares between the actual value, y_i, and the predicted value given by the equation of the line, $\hat{y}_i$. Thus we are trying to minimize: $\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2$.

We're going to skip the proof for what the values for b and m that minimize this equation are since they require the use of calculus, and this section is on the equations of lines. Let $\bar{x}$, pronounced x-bar, represent the mean (or average) x value of all the data points. Respectively $\bar{y}$, pronounced y-bar, is the mean (or average) y value of all the data points.

$$m = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j}{\sum_{i=1}^{n} (x_i^2) - \frac{1}{n} (\sum_{i=1}^{n} x_i)^2}$$

$$b = \frac{1}{n} \sum_{i=1}^{n} y_i - m \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{y} - m\bar{x}$$

Using these values of m and b we now have a line that approximates the points on the graph.

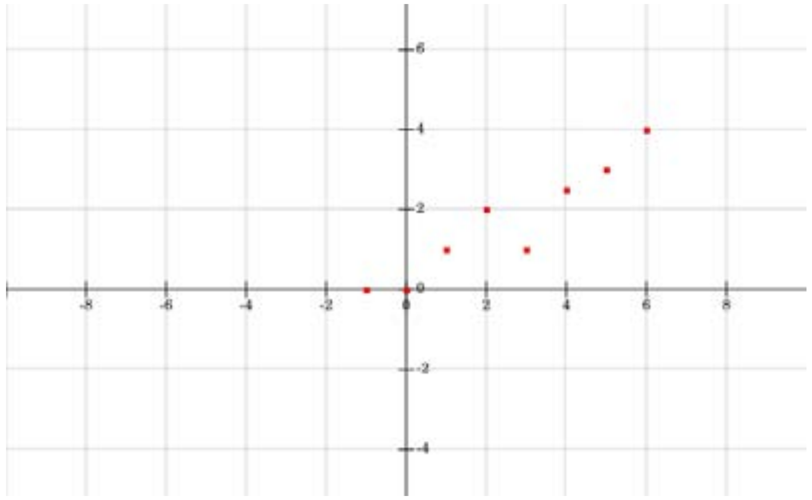


Figure 2.26
Example Points
Here are the points used in the example worked out in the section. The points are (-1, 0), (0, 0), (1, 1), (2, 2), (3, 1), (4, 2.5), (5, 3), (6, 4).

Example

For this example we will be using n=8 points, which are (-1, 0), (0, 0), (1, 1), (2, 2), (3, 1), (4, 2.5), (5, 3), and (6, 4) (*Figure 2.26*). Using these 8 points, we'll plug them into our equations and find the slope and intercept that best approximate this data. First we find the slope, remember:

- $m = \frac{\sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_j}{\sum_{i=1}^{n} (x_i^2) - \frac{1}{n} (\sum_{i=1}^{n} x_i)^2}$
- $\sum_{i=1}^{n} x_i y_i = 0 + 0 + 1 + 4 + 3 + 10 + 15 + 24 = 57$
- $\sum_{i=1}^{n} x_i = -1 + 0 + 1 + 2 + 3 + 4 + 5 + 6 = 20$

- $\sum_{i=1}^{n} y_i = 0 + 0 + 1 + 2 + 1 + 2.5 + 3 + 4 = 13.5$

Therefore the numerator in the slope equation is

$57 - \frac{1}{8}(20)(13.5) = 23.25.$

$$\sum_{i=1}^{n} (x_i^2) = 1 + 0 + 1 + 4 + 9 + 16 + 25 + 36 = 92$$

So the denominator is $92 - \frac{1}{8}(20)^2 = 92 - 50 = 42$ and the slope is

$\frac{23.25}{42} \approx 0.554.$

Now for the y-intercept, b, we have $\bar{x} = \frac{20}{8} = 2.5$ and

$\bar{y} = \frac{13.5}{8} = 1.6875$. Therefore $b \approx 1.6875 - 0.554(2.5) = 0.3025.$

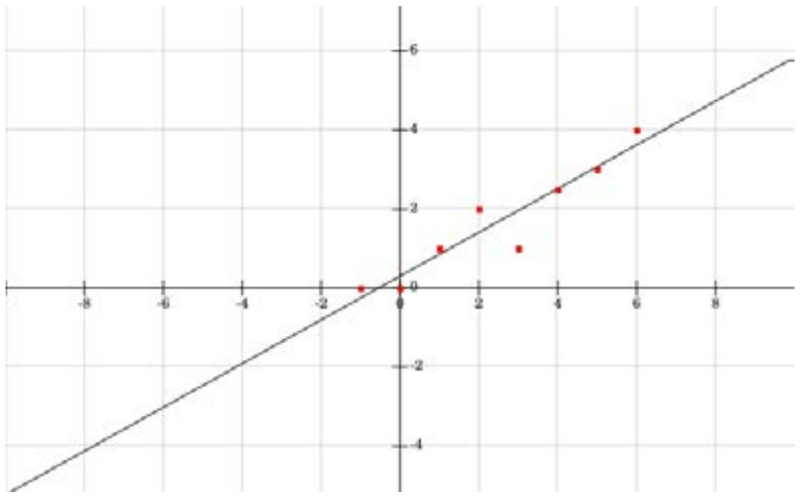


Figure 2.27 Least Squares Fit Line

Here is a line used by the least squares approximation. The line equation is y = 0.554x+0.3025.

Our final equation is therefore $y = 0.554x + 0.3025$, and this line is graphed, along with the points, in *Figure 2.27*. As shown, the approximation is quite good.

Least squares can sometimes fail, for instance if we have a point that is far away from the approximating line, then it will skew the results and make the line much worse. For instance, let's say in our original example, instead of the point (-1, 0), we have (-1, 6). We won't go through all the calculations again, but simply state that if one does the same thing as above with the new point, it ends up as $m \approx 0.0536$ and $b \approx 2.3035$, to get the new equation $y = 0.0536x + 2.3035$, shown in *Figure 2.28*. As can be seen, this new line does not fit the data well due to the **outlier**. Indeed, trying to fit linear models to data that is quadratic, cubic, or anything non-linear, or data with many outliers and errors can result in bad approximations.

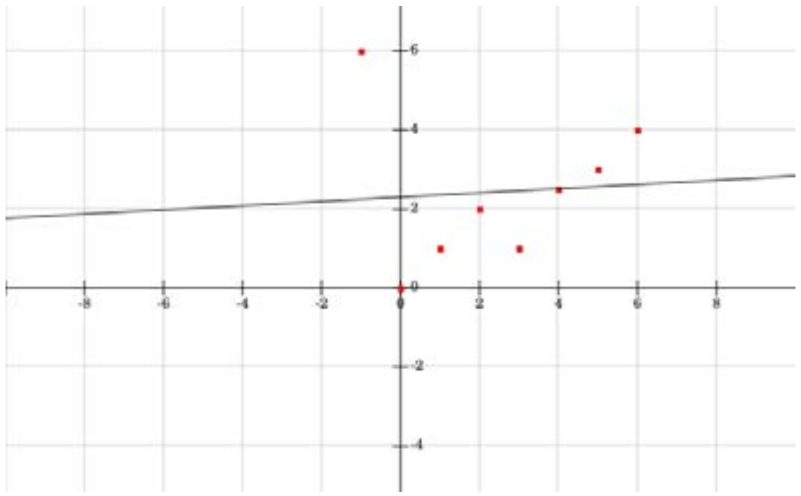


Figure 2.28 Outlier Approximated Line

Here is the approximated line given the new outlier point at (-1, 6).

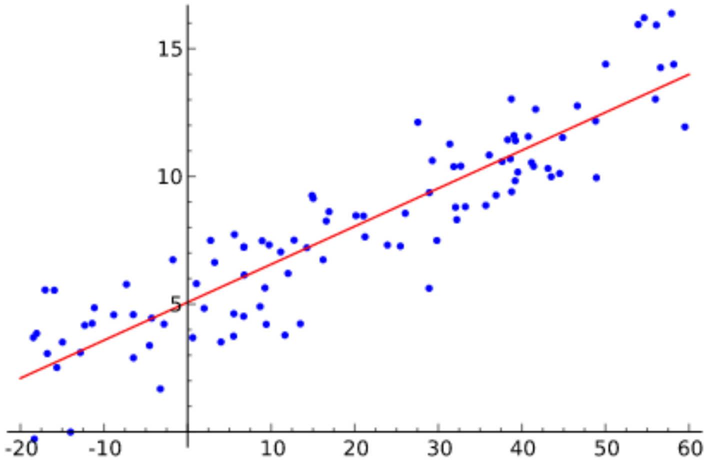


Figure 2.29 Line Fitting

Fitting a line to a curve using linear regression.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/modeling-equations-of-lines/fitting-a-curve/

Section 4

Functions Revisited

Increasing, Decreasing, and Constant Functions

Relative Minimums and Maximums

Piecewise Functions

https://www.boundless.com/algebra/graphs-functions-and-models/functions-revisited/

Increasing, Decreasing, and Constant Functions

Functions can either be constant with a single value, increasing as x increases, or decreasing as x increases.

KEY POINTS

- A constant function is a function whose values do not vary, regardless of the input into the function.
- An increasing function is one where for every x1 and x2 that satisfies x2 > x1, then $f(x_2) \geq f(x_1)$. If it is strictly greater than, then it is strictly increasing.
- A decreasing function is one where for every x1 and x2 that satisfies x2 > x1, then $f(x_2) \leq f(x_1)$. If it is strictly less than, then it is strictly decreasing.

Constant Functions

In mathematics, a **constant function** is a function whose values do not vary—they are constant. For example, the function f(x) = 4 (*Figure 2.30*) is constant since f maps any value to 4. More formally, a function f : A → B is a constant function if f(x) = f(y) for all values of x and y in A. Every empty function is constant, vacuously, since there are no values of x and y in A for which f(x) and f(y) are different when A is the empty set. In polynomial functions, a non-zero constant function is called a polynomial of degree zero. A function is said to be "identically zero" if it takes the value 0 for every argument; it is trivially a constant function.

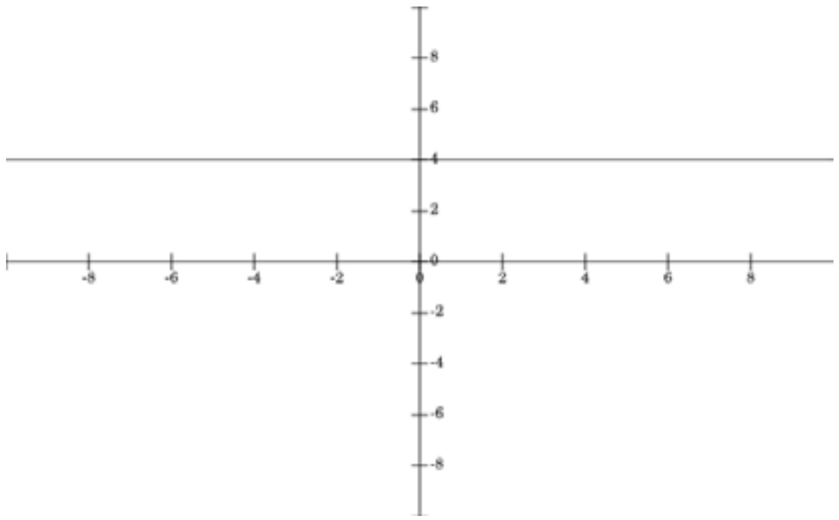


Figure 2.30
Constant Function
A constant function f(x)=4

A **composite function** is an application of one function to the results of another. Composing two functions—f(x) and g(x) for example—is written as either f(g(x)) or:

$$f \circ g(x)$$

We won't go into details in this section, but say you have two functions:

$$f(x) = 2x - 2 \ and \ g(x) = 3x - 1$$

The composite f(g(x)) would be:

$f(g(x)) = f \circ g(x) = f(3x - 1) = 2(3x - 1) - 2$

Several key points about constant functions with respect to composite functions are:

- f : A → B is a constant function.
- For all functions g, h : C → A, $f \circ g = f \circ h$.
- The composition of f with any other function is also a constant function.

Increasing and Decreasing Functions

An **increasing function** is a function such for every x_1 and x_2 that satisfies $x_2 > x_1$, then $f(x_2) \geq f(x_1)$, if the equality is strictly greater than (not greater than or equal to), then the function is "strictly increasing." A **decreasing function** is a function such that for every x_1 and x_2 that satisfies $x_2 > x_1$, then $f(x_2) \leq f(x_1)$. If the equality is strictly less than (not less than or equal to), then the function is "strictly decreasing."

In terms of the linear function $f(x) = mx + b$, if m is positive, the function is increasing, if m is negative, it is decreasing, and if m is zero, the function is a constant function.

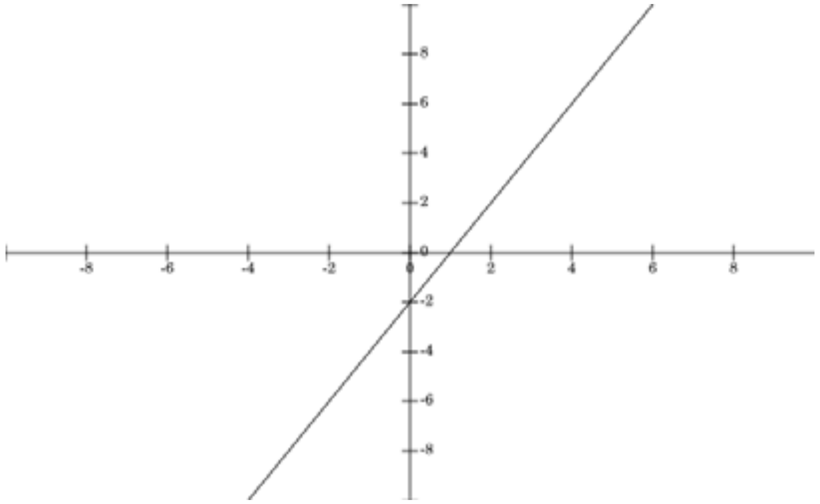


Figure 2.31 Increasing Linear Function

This is an increasing linear function, f(x)=2x-2.

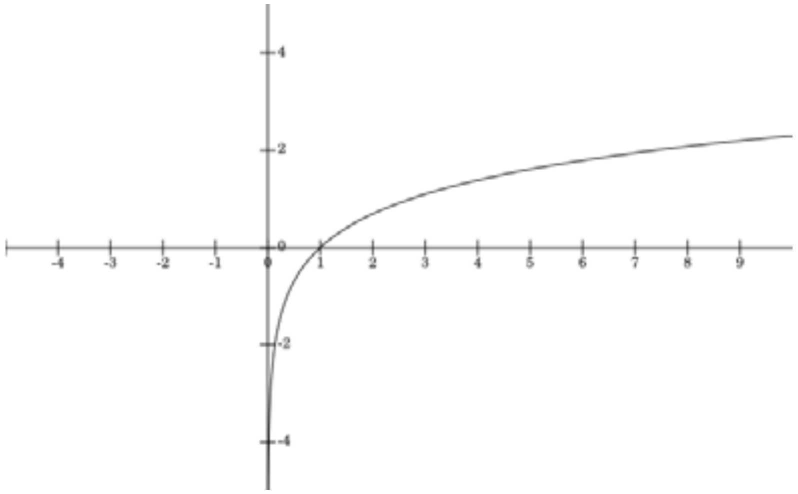


Figure 2.33 Natural Logarithm

This is the natural logarithm, or ln(x), and is an increasing function.

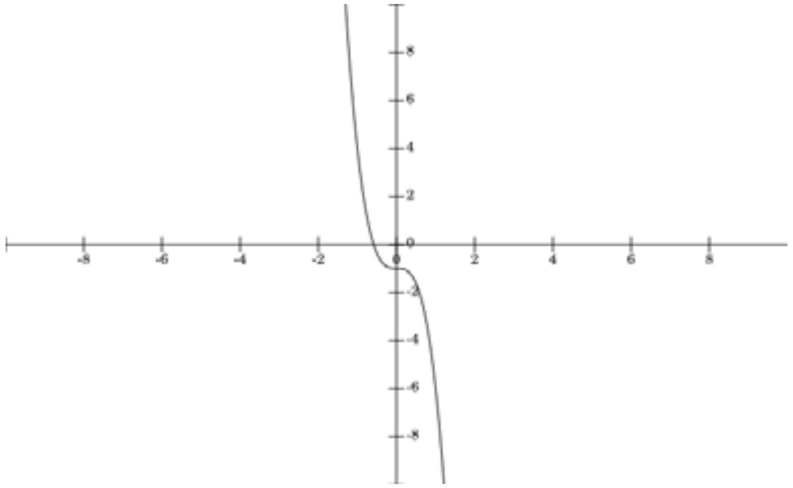


Figure 2.32 A Cubic Function

This is the function $f(x) = -5x^3 - 1$, and is a decreasing function.

Relative Minimums and Maximums

Relative minima and maxima are points of the smallest and greatest values in their neighborhoods respectively.

KEY POINTS

- Minima and maxima are collectively known as extrema.
- A function has a global (or absolute) maximum point at $x*$ if $f(x*) \geq f(x)$ for all x. Similarly, a function has a global (or absolute) minimum point at $x*$ if $f(x*) \leq f(x)$ for all x.
- A real-valued function f defined on a real line is said to have a relative maximum at the point $x*$, if there exists some $\varepsilon > 0$ such that $f(x*) \geq f(x)$ when $|x - x*| < \varepsilon$. Similarly, a function has a relative minimum at $x*$, if $f(x*) \leq f(x)$ when $|x - x*| < \varepsilon$.
- Functions don't necessarily have extrema in them. For example any line, $f(x) = mx+b$ where m and b are constants, does not have any extrema, be they local or global.

In mathematics, the **maximum** and **minimum** of a function, known collectively as **extrema**, are the largest and smallest value that a function takes at a point either within a given neighborhood (local or relative extremum) or within the function domain in its entirety (global or absolute extremum).

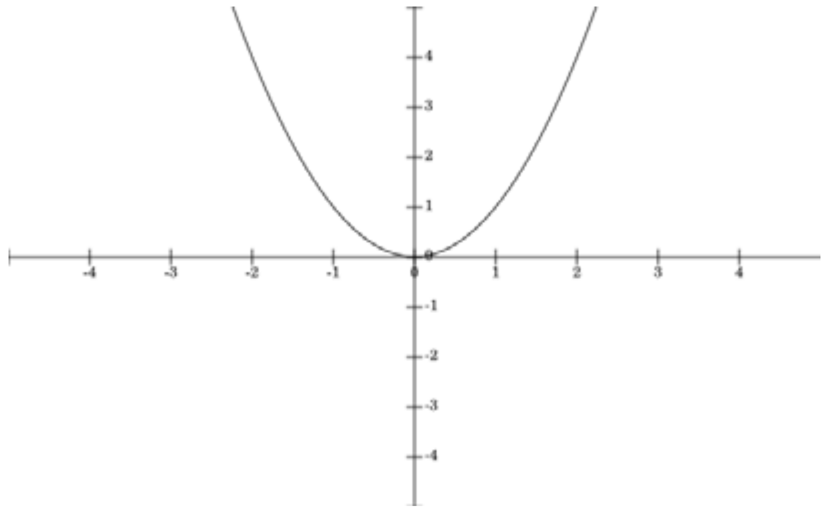


Figure 2.34 Global and Relative Minimum Example

This function, f(x) = x2, has a global (and also relative) minimum at x=0.

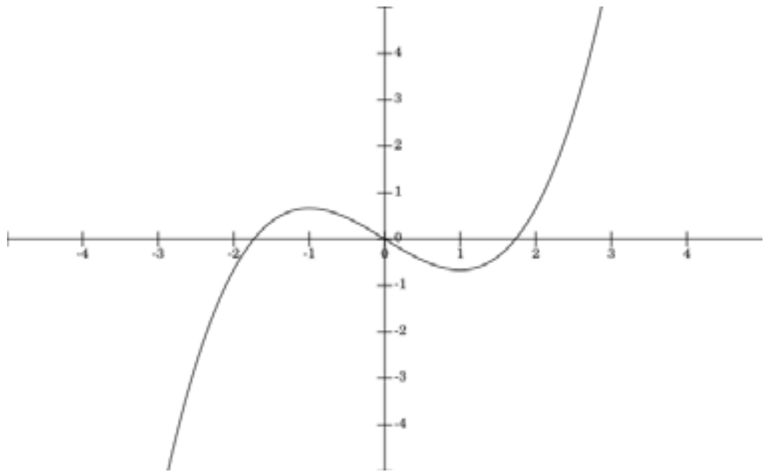


Figure 2.35 Relative Minimum and Maximum Example

In this function, f(x) = (x3)/3-x, there is a relative maximum at x=-1 and a relative minimum at x=1.

A real-valued function, f, defined on a real line, is said to have a relative maximum point at the point x∗, if there exists some ε > 0 such that f(x∗) ≥ f(x) when |x − x∗| < ε. The value of the function at this point is called the maximum of the function. Similarly, a function has a relative minimum point at x∗, if f(x∗) ≤ f(x) when |x − x∗| < ε. The value of the function at this point is called the minimum of the function. That is to say, a point is a relative maximum if there is no point an infinitesimally small distance away to the left or right that is greater than this point. Similarly, a point is a relative minimum if there is no point an infinitesimally small distance away to the left or right that is less than this point.

A function has a global (or absolute) maximum point at x∗ if f(x∗) ≥ f(x) for all x. Similarly, a function has a global (or absolute) minimum point at x∗ if f(x∗) ≤ f(x) for all x. Global extrema are also relative extrema.

Functions may not have any extrema in them, such as the line y = x. This line increases towards infinite and decreases towards negative infinite, and has no relative extrema.

Why should we care about minima and maxima? They're used heavily in optimization problems and artificial intelligence where,

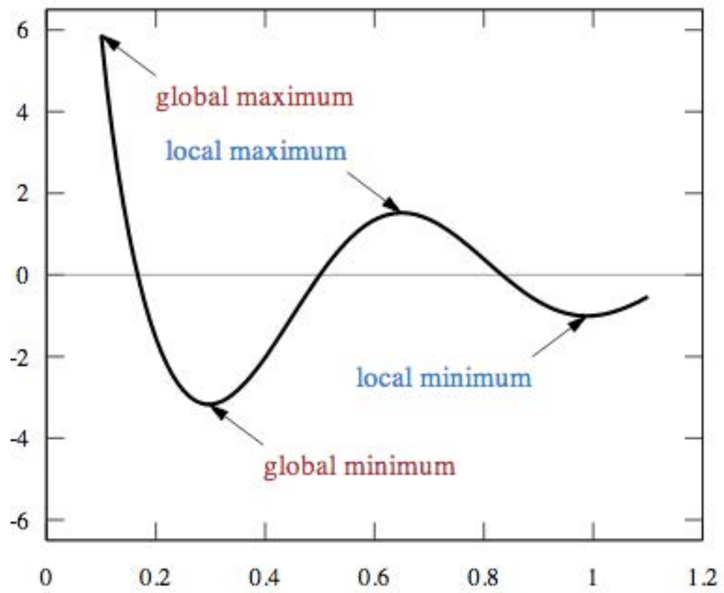


Figure 2.36 Examples of Relative and Global Extrema

This graph has examples of all four possibilities: relative maximum and minimum, and global maximum and minimum.

given a number of constraints on resources, we want the best use of our resources. For instance, we may want to maximize our profits given the items we can make and our resources. In artificial intelligence, we may want to discover what the least costly plan of action to take is for a robot. Ideally you'd want to find the global minima for the plans. However, because there is not unlimited time to identify the right plan, artificial intelligence often simply finds the local minima.

EXAMPLE

The function $f(x) = x^2$ has a global (and relative) minimum at x=0.

The function $f(x) = \frac{x^3}{3} - x$ has a local maximum at x=-1 and a local minimum at x=1.

The aforementioned minima and maxima can be discovered using calculus and identifying where the first derivative crosses 0. However, because this is algebra and not calculus, we will not describe how to find a maxima and minima. That being said, it is important to know what they are and how they are applied.

Piecewise Functions

A piecewise function is a function defined by multiple subfunctions that are each applied to separate intervals of the input.

KEY POINTS

- Piecewise functions are defined using the common functional notation, where the body of the function is an array of functions and associated subdomains.
- The absolute value |x| is a very common piecewise function. For a real number, its value is -x when x < 0 and x when $x \geq 0$.
- Piecewise functions may have horizontal or vertical gaps (or both) in their functions. A horizontal gap means that the function is not defined for those inputs.
- An open circle at the end of an interval in one of the subdomains means that the end point is not included in the interval, i.e. strictly less than or strictly greater than. A closed circle means the end point is included.

In mathematics, a piecewise-defined function, also called a **piecewise function**, is a function which is defined by multiple subfunctions, each subfunction applying to a certain interval of the main function's domain, a **subdomain**. Piecewise is actually a way of expressing the function, rather than a characteristic of the function itself, but with additional qualification, it can describe the nature of the function.

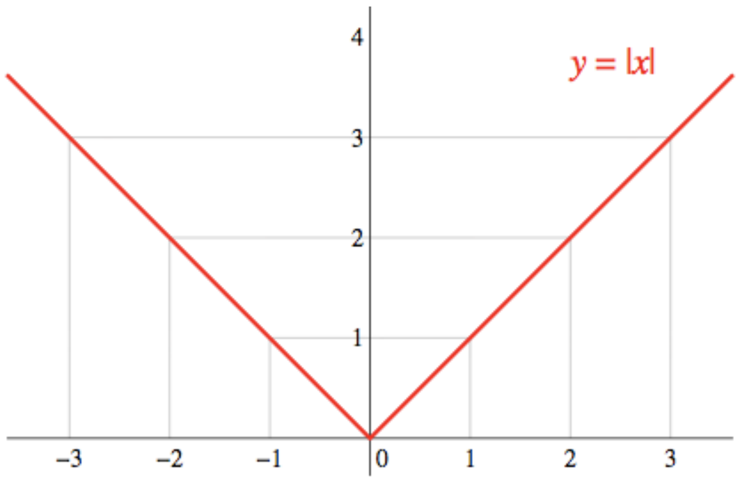


Figure 2.38 **Absolute Value**

The absolute value is a piecewise function where f(x) = x when x >= 0 and f(x) = -x when x < 0.

Notation and Interpretation

Piecewise functions are defined using the common functional notation, where the body of the function is an array of functions and associated subdomains. Crucially, in most settings, there must only

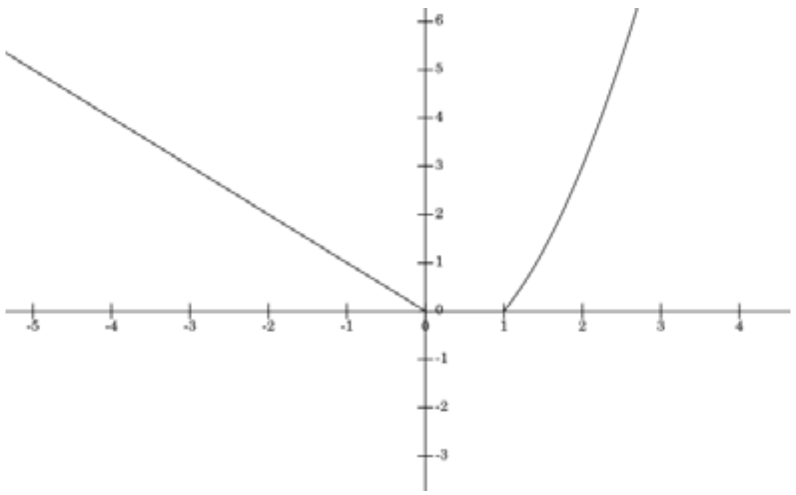


Figure 2.37 Gap in the Input

For this piecewise function that defines f(x) = -1 when x <= 0, and f(x) = x^2 when x >= 1.

be a finite number of subdomains, each of which must be an interval, in order for the overall function to be called "piecewise". For example, consider the piecewise definition of the **absolute value** function:

$$|x| = \begin{cases} -x, & if\ x < 0 \\ x, & if\ x \geq 0 \end{cases}$$

For all values less than zero, the first function (−x) is used, which negates the sign of the input value, making negative numbers positive. For all values of x greater than or equal to zero, the second function (x) is used, which evaluates trivially to the input value itself (*Figure 2.38*).

Jumps

Piecewise functions can have jumps in either the input or the output.

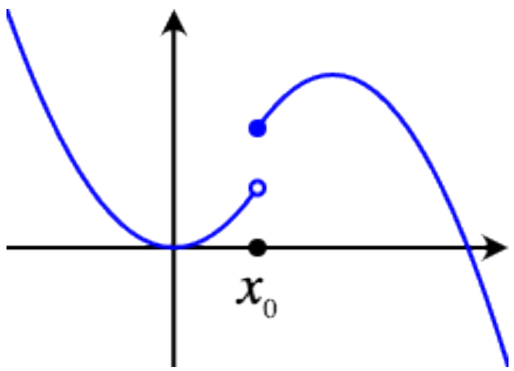


Figure 2.39 Vertical Jump of a Piecewise Function

There is a vertical jump in the piecewise function at x=1.

If there are gaps in the input, then the function is not defined over those input values. For example, the piecewise function: $f(x) = \begin{cases} -x, & if\ x \leq 0 \\ x^2, & if\ x \geq 1 \end{cases}$, x is not defined between 0 and 1, and there is a noticeable visible gap between the two parts of the function *Figure 2.37*. Therefore the domain of this function is $x \leq 0\ and\ x \geq 1$.

Piecewise functions can also have vertical gaps. An example of this is the piecewise function: $f(x) = \begin{cases} x^2, & if\ x < 1 \\ -(x-2)^2+3, & if\ x \geq 1 \end{cases}$, has a vertical jump at x = 1. Since $1^2 = 1$ and $-(1-2)^2+3 = 2$, not the open circle as on the first equation, x^2. The open circle means that it does not include that value, this is because the domain on the first function is x < 1, strictly less than, therefore x does not equal 1. However, x can equal 0.9, or 0.99, or 0.9999, and so on, so there has to be some representation that it goes up to 1, but never actually equals 1. A closed circle means that point is included, and the second part of this equation does have a closed circle associated with it.

Section 5

Algebra of Functions

https://www.boundless.com/algebra/graphs-functions-and-models/algebra-of-functions/

Sums, Differences, Products, and Quotients

Adding, subtracting, multiplying, and dividing equations requires one to follow some key steps that can simplify the problem.

KEY POINTS

- Adding and subtracting equations involves grouping like terms and carrying out basic arithmetic.
- Multiplying and dividing monomials applies to both like and unalike terms.
- Multiplying binomials and trinomials follows the FOIL method.

Adding and Subtracting Functions

Terms in algebraic equations are separated by + or –. For instance, in the equation y = x + 5, there are two terms, while in the equation y = $2x^2$, there is only one term.

In adding equations, it is important to collect like terms to simplify the expression. "Like terms" are those that have the same kind of variable. For instance, take two equations:

$f(x) = x + 5$ and $g(x) = 2x - 3$

By adding these two equations together, we get:

$f(x) + g(x) = h(x) = x + 5 + 2x - 3$

We then collect like terms. In this case, "x" and "2x" are like terms, as are "5" and "-3." The result is:

$h(x) = 3x + 2$

Subtracting two equations follows similar logic, except a negative sign is applied to an entire equation. In this case, the -3 of g(x) becomes +3.

$f(x) - g(x) = h(x) = x + 5 - (2x - 3) = x + 5 - 2x + 3$

$h(x) = -x + 8$

It is important to remember to only add together like terms. For instance, x^2 cannot be added to x because they are not the same term, although they use the same variable.

Multiplying and Dividing Functions

Adding and subtracting functions is quite straightforward, as shown. While adding and subtracting equations only affects like terms, multiplying and dividing functions affects all terms equally.

It is easiest to start with monomials. A **monomial** equations has one term; a **binomial** has two terms; a trinomial has three terms.

$f(x) = 3$ and $g(x) = x$

$f(x) * g(x) = 3 * x = 3x$

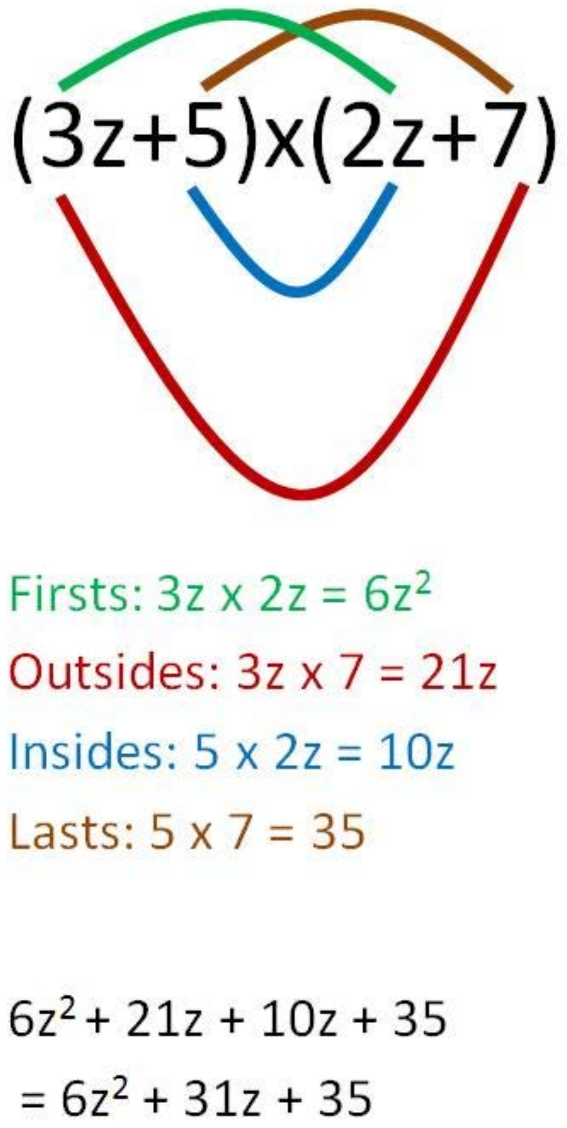


Figure 2.40 FOIL Method Diagram

A color-coded diagram of the FOIL method using both numbers and variables.

In this case, two unlike terms (3 and x) could be multiplied. If the terms both contain the same variables, their exponents are added together and their multipliers are multiplied.

$f(x) = 3x$ and $g(x) = 2x^3$

$f(x) * g(x) = (3x) * (2x^3) = 6x^4$

When there are multiple variables, the two exponents are added separately and the variables remain next to each other.

$f(x) = -8ab$ and $g(x) = 9a^3b$

$f(x) * g(x) = -36a^4b^2$

Multiplying binomials and trinomials is more complicated, and follows the FOIL method. FOIL is a mnemonic for the standard method of multiplying two binomials; the method may be referred to as the FOIL method. FOIL is an acronym for the four terms of the product:

- First (“first” terms of each binomial are multiplied together)
- Outer (“outside” terms are multiplied—that is, the first term of the first binomial and the second term of the second)
- Inner (“inside” terms are multiplied—second term of the first binomial and first term of the second)

- Last ("last" terms of each binomial are multiplied)

The general form is shown in [Figure 2.40](#) and is diagrammed in [Figure 2.41](#).

Dividing equations uses similar theory as multiplying, since division is the equivalent of multiplying by the inverse.

Figure 2.41 General Form of the FOIL Method

$$(a+b)(c+d) = \underbrace{ac}_{\text{first}} + \underbrace{ad}_{\text{outside}} + \underbrace{bc}_{\text{inside}} + \underbrace{bd}_{\text{last}}$$

The general form of the FOIL method using only variables as the potential multipliers.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/algebra-of-functions/sums-differences-products-and-quotients/

Difference Quotients

The difference quotient is used in algebra to calculate the average slope between two points but has broader effects in calculus.

KEY POINTS

- Two points on a graph may give different results when plugged into a function. The difference between two distinct points is known as their Delta (Δ) P, as is the difference in their function result.
- The function difference between two point values divided by the point difference is known as the difference quotient.
- The difference quotient is essentially the average slope of a function between two points.
- In calculus, the difference quotient is used to calculate the derivative when the difference between the two points is infinitesimally small.

The primary vehicle of calculus and other higher mathematics is the function. Its "input value" is its argument, usually a point ("P") expressible on a graph. The difference between two distinct points, themselves, is known as their Delta (Δ) P, as is the difference in their function result, the particular notation being determined by the direction of formation:

- Forward difference: $\Delta F(P) = F(P + \Delta P) - F(P)$
- Central difference: $\delta F(P) = F(P + 1/2\Delta P) - F(P - 1/2\Delta P)$
- Backward difference: $\nabla F(P) = F(P) - F(P - \Delta P)$

The function difference divided by the point difference is known as the difference quotient, attributed to Isaac Newton. It is also known as Newton's quotient:

$$\frac{\Delta F(P)}{\Delta P} = F(P + \Delta P) - \frac{F(P)}{\Delta P} = [\nabla F(P + \Delta P)]\Delta P$$

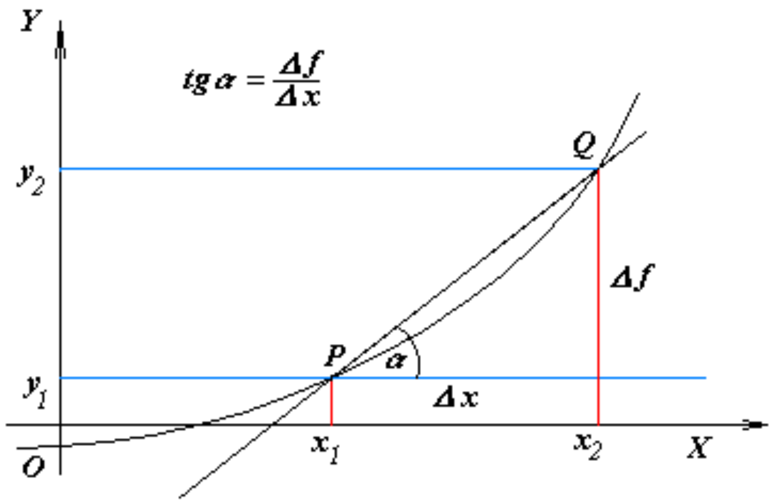


Figure 2.42 Difference Quotient Chart

The difference quotient can be used to calculate the average slope (here, represented by a straight line) between two points P and Q.

The difference quotient is the average slope of a function between two points. However, it is important to note it is not necessarily to actual slope of the curve, as can be visually seen in *Figure 2.42*. To show how the above equation can be written as the average slope, put it into more familiar terms. The points can be represented represent as (P, F(P)) and (P+ΔP, F(P+ΔP)) as (x, f(x)) and (x+h, f(x +h)).

Remember, the formula for slope is

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

This becomes:

$$m = \frac{f(x + h) - f(x)}{(x + h) - x}$$

Which simplifies to:

$$m = \frac{f(x + h) - f(x)}{h}$$

If you look carefully, this formula is the same as $\frac{F(P + \Delta P) - F(P)}{\Delta P}$, just written differently.

Applications in Calculus

If |ΔP| is finite, meaning measurable, then ΔF(P) is known as a finite difference, with specific denotations of DP and DF(P). This is the case for all algebraic applications.

If |ΔP| is infinitesimal, an infinitely small amount usually expressed in standard analysis as a limit, then ΔF(P) is known as an

infinitesimal difference, with specific denotations of dP and dF(P). In this case, the difference quotient is know as a **derivative**, a useful tool in calculus.

Composition of Functions and Decomposing a Function

Functional composition allows for the application of one function to another; this step can be undone by using functional decomposition.

KEY POINTS

- Functional composition applies one function to the results of another.
- Functional decomposition resolves a functional relationship into its constituent parts so that the original function can be reconstructed from those parts by function composition.
- Decomposition of a function into non-interacting components generally permits more economical representations of the function.

Function Composition

Function composition is the application of one function to the results of another. For instance, the functions $f: X \rightarrow Y$ and $g: Y \rightarrow Z$ can be composed by computing the output of g when it has an argument of f(x) instead of x. Intuitively, if z is a function g(y) and y is a function f(x), then z is a function of x.

Thus, one obtains a composite function $g \circ f: X \to Z$ defined by $(g \circ f)(x) = g(f(x))$ for all x in X. The notation $g \circ f$ is read as "g circle f", or "g composed with f", "g after f", "g following f", or just "g of f".

The composition of functions is always associative. That is, if f, g, and h are three functions with suitably chosen domains and codomains, then $f \circ (g \circ h) = (f \circ g) \circ h$, where the parentheses serve to indicate that composition is to be performed first for the parenthesized functions. Since there is no distinction between the choices of placement of parentheses, they may be safely left off.

Functional Decomposition

Functional decomposition broadly refers to the process of resolving a functional relationship into its constituent parts in such a way that the original function can be reconstructed (i.e., recomposed) from those parts by function composition. In general, this process of decomposition is undertaken either for the purpose of gaining insight into the identity of the constituent components (which may reflect individual physical processes of interest), or for the purpose of obtaining a compressed representation of the global function; a task which is feasible only when the constituent processes possess a certain level of modularity (i.e., independence or non-interaction).

For a multivariate function $y = f(x_1, x_2, \ldots, x_n)$, functional decomposition generally refers to a process of identifying a set of functions $\{ g_1, g_2, \ldots, g_m \}$ such that

$$f(x_1, x_2, \ldots, x_n) = \phi(g_1(x_1, x_2, \ldots, x_n), g_2(x_1, x_2, \ldots, x_n), \ldots, g_m(x_1, x_2, \ldots, x_n))$$

where φ is some other function. Thus, we would say that the function f is decomposed into functions $\{g_1, g_2, \ldots, g_m\}$. As illustrated in *Figure 2.43*, this process is intrinsically hierarchical in the sense that we can (and often do) seek to further decompose the functions into a collection of constituent functions $\{ h_1, h_2, \ldots, h_p \}$ such that

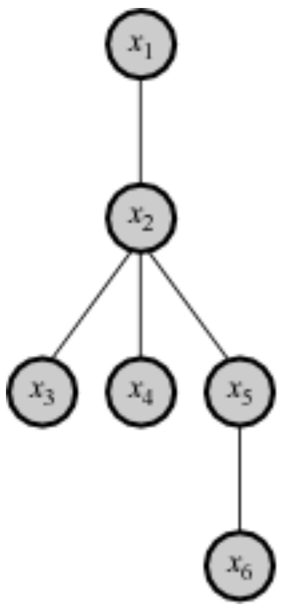


Figure 2.43 A Chow-Liu Tree

An example of a sparsely connected dependency structure. Direction of causal flow is upward.

$$g_i(x_1, x_2, \ldots, x_n) = \gamma(h_1(x_1, x_2, \ldots, x_n), h_2(x_1, x_2, \ldots, x_n), \ldots, h_p(x_1, x_2, \ldots, x_n))$$

In general, functional decompositions are worthwhile when there is a certain "sparseness" in the dependency structure i.e. when constituent functions are found to depend on approximately disjointed sets of variables. Also, decomposition of a function into

non-interacting components generally permits more economical representations of the function.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/algebra-of-functions/composition-of-funcitons-and-decomposing-a-function/

Section 6

Transformations

Symmetry

Even and Odd Functions

Transformations of Functions

Translations

Reflections

Stretching and Shrinking

https://www.boundless.com/algebra/graphs-functions-and-models/transformations/

Symmetry

Two objects are symmetric if one object is obtained from the other by a transformation.

KEY POINTS

- If one object is obtained from the other by an invariant transformations, the two objects are symmetric to each other.
- The value of the output is invariant under permutations of variables in the case of symmetric functions.
- A binary operation is commutative if the operator, as a function of two variables, is a symmetric function.

Two objects are symmetric to each other with respect to the invariant transformations if one object is obtained from the other by one of the transformations. It is an equivalence relation. In the case of symmetric functions, the value of the output is invariant under permutations of variables. These permutations form the symmetric group.

Symmetric Functions

From the form of an equation, one may observe that certain permutations of the unknowns result in an equivalent equation. In that case, the set of solutions is invariant under any permutation of the unknowns in the group generated by the aforementioned permutations. For example:

$(a-b)(b-c)(c-a)=10$; for any solution (a,b,c), permutations (a b c) and (a c b) can be applied, giving additional solutions (b, c, a) and (c, a, b). $a^2c+3ab+b^2c$ remains unchanged under interchanging of a and b.

For a sphere, if φ is the longitude, θ the colatitude, and r the radius, as illustrated in *Figure 2.44*, then the great-circle distance is given by:

$$d(\theta_1,\phi_1,\theta_2,\phi_2)=r\cos^{-1}(\cos\theta_1\cos\theta_2+\sin\theta_1\sin\theta_2\cos(\phi_1-\phi_2)).$$

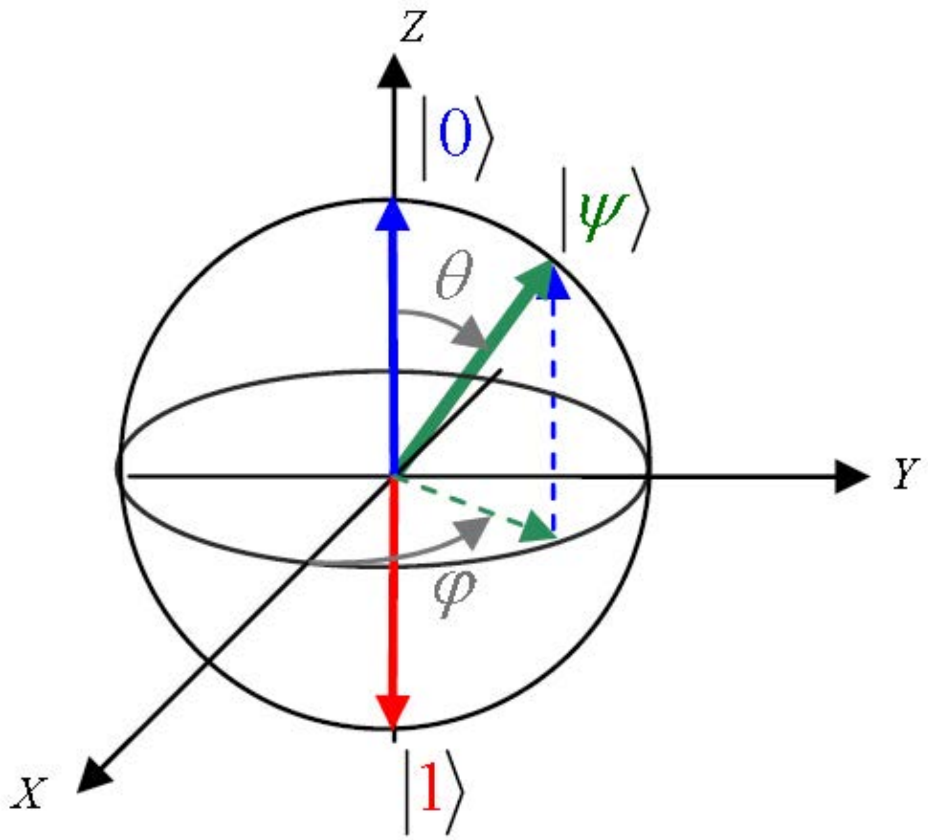
Some symmetries cleared from the problem can be verified in the formula; the distance is invariant under:

- adding the same angle to both longitudes
- interchanging longitudes and/or interchanging latitudes
- reflecting both colatitudes in the value 90°

In Algebra

A relation is symmetric if, and only if, the corresponding **boolean-valued function** is a symmetric function. A binary operation is **commutative** if the operator, as a function of two variables, is a

Figure 2.44 Sphere



In this sphere, ϕ is the longitude and θ the colatitude. The radius is the green dotted line. The notation of |0〉, |1〉, and |ψ〉 may be unfamiliar (this is Bra-Ket notation), but none of the lengths depicted by 0, 1 and ψ is essential to the discussion of symmetry.

symmetric function. Symmetric operators on sets include the union, intersection, and symmetric difference.

A symmetric matrix, seen as a symmetric function of the row and column number, is an example of a symmetric function in algebra. Another example is the second order partial derivatives of a smooth function, seen as a function of the two indexes.

Even and Odd Functions

Functions that have an additive inverse can be classified as odd or even depending on their symmetry properties.

KEY POINTS

- The parity of a function does not necessarily reveal whether the function is odd or even.
- An even function is symmetric about the y-axis.
- An odd function is symmetric with respect to 180-degree rotation about the origin.

Functions can be classified as "odd" or "even" based on their composition. These labels correlate with symmetry properties of the function.

The terms "odd" and "even" can only be applied to a limited set of functions. For a function to be classified as one or the other, it must have an **additive inverse**. Therefore, it must have a number that, when added to it, equals 0.

Oftentimes, the **parity** of a function will reveal whether it is odd or even. For example, the function $f(x)=x^2$ is even and has an exponent (2) that is an even integer. This does not apply in every instance, however. For example, $f(x)=|x|$ has an exponent that is of an odd integer, but is also an even function.

Even Functions

Even functions are algebraically defined as functions such that the following relationship holds for all values of x:

$$f(x) = f(-x)$$

Functions that satisfy the requirements of being even are symmetric about the y-axis. Therefore, a reflection about the y-axis produces no change in the points on the graph.

Examples of even functions include: x^2, x^4, $|x|$, and $\cos(x)$ (Figure 2.45).

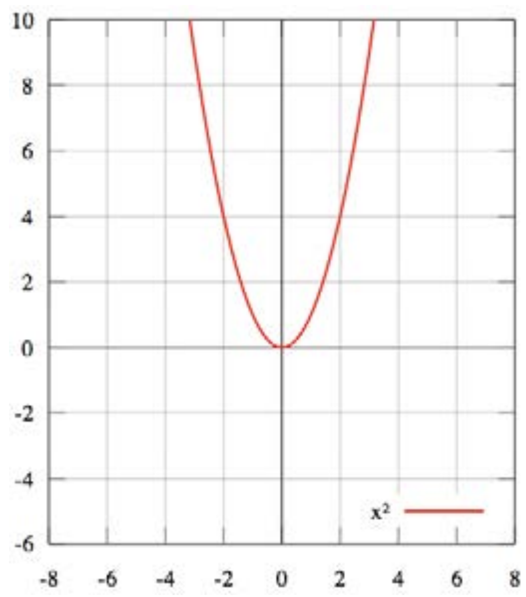


Figure 2.45 Graph of y=x^2

y=x^2 is an even function; the order of x is even (2), and it is symmetrical with respect to the vertical axis.

Odd Functions

Odd functions are algebraically defined as functions such that the following relationship holds for all values of x:

$$-f(x) = f(-x)$$

This relationship can also be expressed as:

$$f(x) + f(-x) = 0$$

Functions that satisfy the requirements of being odd are symmetric with respect to rotation about the origin. In other words, rotating the graph 180 degrees about the point of origin results in the same, unchanged graph.

Examples of odd functions include: x, x3, x5, and sin(x) (*Figure 2.46*).

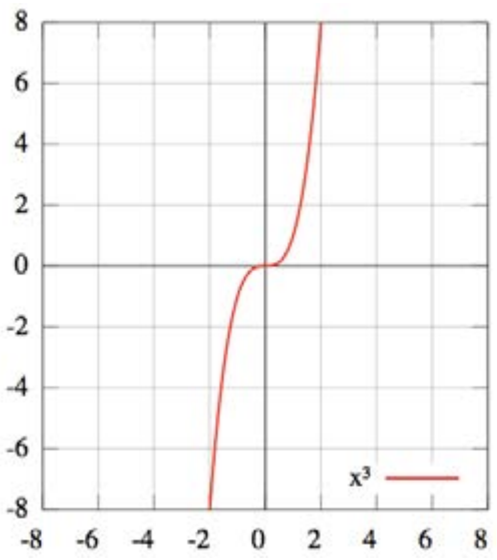


Figure 2.46 Graph of y=x^3

y=x^3 is an odd function; the order of x is odd (3), and it is symmetrical when rotated 180 degrees about the origin.

Transformations of Functions

Transformations alter a function while maintaining the original characteristics of that funcction.

KEY POINTS

- Transformations are ways that a function can be adjusted to create new functions.
- Transformations often preserve the original shape of the function.
- Common types of transformations include rotations, translations, reflections, and scaling (also known as stretching/shrinking).

A transformation could be any function mapping a set, X, on to another set or on to itself. However, often the set X has some additional algebraic or geometric structure and the term "transformation" refers to a function from X to itself that preserves this structure.

Examples include translations, reflections, rotations, and scaling. These can be carried out in **Euclidean space**, particularly in dimensions 2 and 3. They are also operations that can be performed using linear algebra and described explicitly using matrices.

A translation, or translation operator, is an affine transformation of Euclidean space which moves every point by a fixed distance in the same direction. It can also be interpreted as the addition of a constant vector to every point, or as the shifting of the origin of the coordinate system. In other words, if v is a fixed vector, then the translation Tv will work as $T_v(p) = p + v$. A graphical representation of vertical translations can be viewed in *Figure 2.47*.

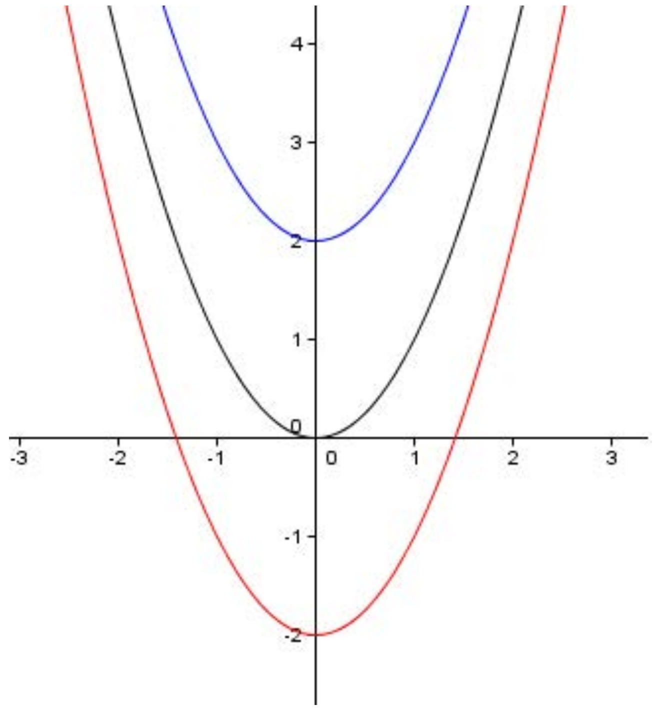


Figure 2.47 Basic Quadratic Equation

The black curve indicates the original function while the red and blue curves represent two vertical translations of that function.

A reflection is a map that transforms an object into its mirror image. In geometry a "mirror" is a hyperplane of fixed points. For example, a reflection of the small English letter p in respect to a vertical line would look like q. In order to reflect a planar figure one needs the "mirror" to be a line (axis of reflection or axis of

symmetry), while for reflections in the three-dimensional space one would use a plane (the plane of reflection or symmetry) for a mirror.

A rotation is a transformation that is performed by "spinning" the object around a fixed point known as the center of rotation. You can rotate your object at any degree measure but 90° and 180° are two of the most common.

Uniform scaling is a linear transformation that enlarges or diminishes objects. The scale factor is the same in all directions; it is also called a homothety or dilation. The result of uniform scaling is similar (in the geometric sense) to the original. Scaling can also be referred to as "stretching" or "shrinking" a function.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/transformations/transformations-of-functions/

Translations

A translation of a function is a shift in one of the cardinal directions; it is represented by adding or subtracting from either y or x.

KEY POINTS

- A translation is a function that moves every point a constant distance in a specified direction.
- A vertical translation is generally given by the equation $y = f(x) + b$. These translations shift the whole function up or down the y-axis.
- A horizontal translation is generally given by the equation $y = f(x-a)$. These translations shift the whole function side to side on the x-axis.

A translation moves every point in a function a constant distance in a specified direction. In algebra, this essentially manifests as a vertical or horizontal shift of a function. A translation can be interpreted as the addition of a constant **vector** to every point or as shifting the origin of the coordinate system.

Let's use the basic quadratic function to explore translations. In *Figure 2.48*, we can see three functions. The function in black is the

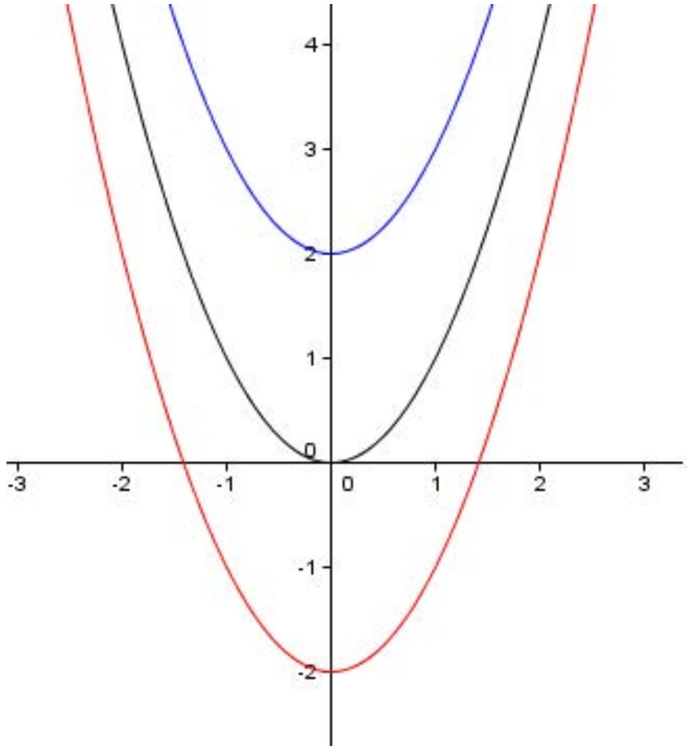


Figure 2.48 Basic Quadratic Equation Vertical Translations

In this function f(x) = x^2, the basic function is black. The blue and red functions are representations of the function translated up or down by two on the y-axis.

untransformed function f(x)=x². The blue function has been translated up by 2. The red function has been translated down by 2.

The equation for the blue line is:

$$y = f(x) + 2 = x^2 + 2$$

The equation for the red line is:

$$y = f(x) - 2 = x^2 - 2$$

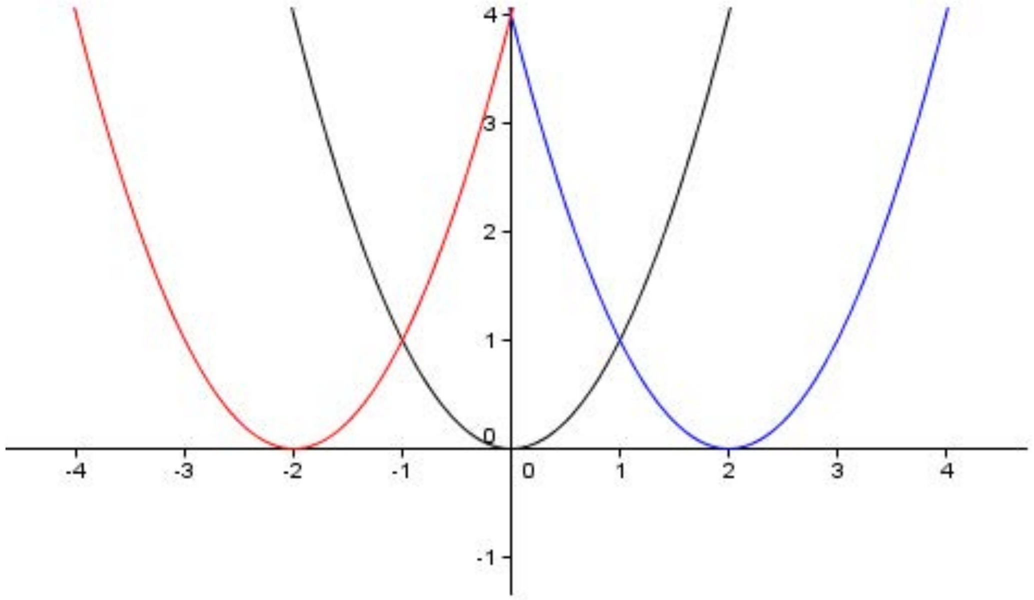
These transformations are fairly straightforward. If a positive number is added, the function shifts up by the same amount. If a positive number is subtracted, the function shifts down by the same amount. In general, a vertical translation is given by the equation:

$$y = f(x) + b$$

Continuing with the basic quadratic function, lets look at horizontal translations as in *Figure 2.49*. Again, the basic black function is $f(x) = x^2$.

The red function has been shifted to the left by 2 and has the equation:

Figure 2.49 Basic Quadratic Equation Horizontal Translations



In this function f(x) = x^2, the basic function is black. The blue and red functions are representations of the function translated up or down by two on the x-axis.

$y = f(x + 2) = (x + 2)^2$

The blue function has been shifted to the left by 2 and has the equation:

$y = f(x - 2) = (x - 2)^2$

The general equation for a horizontal translation is given by:

$y = f(x - a)$

Note that the general form is x–a. When a is positive, the function is shifted to the right. When a is negative the shift is to the left.

Reflections

Reflections are a type of transformation that move an entire curve such that its mirror image lies on the other side of the y- or x-axis.

KEY POINTS

- A reflection swaps all of the x or y values across the y- or x-axis, respectively. It can be visualized by imagining a mirror lies across that axis.
- A vertical reflection is given by the equation y = −f(x) and results in the curve being "reflected" across the x-axis.
- A horizontal reflection is given by the equation y = f(−x) and results in the curve being "reflected" across the y-axis.

Reflections are another type of **transformations** that can be done with functions. A reflection is a mapping from a Euclidean space to itself that is an **isometry** with a hyperplane as a set of fixed points. This set is called the axis (in dimension 2) or plane (in dimension 3) of reflection. The image of a figure by a reflection is its mirror image in the axis or plane of reflection.

In algebra, reflections are generally done over either the x- or y-axis. *Figure 2.50* shows two types of reflections of an exponential function. A vertical reflection is a reflection across the x-axis as

shown below by the red function. The blue function has been reflected horizontally across the y-axis.

A vertical reflection is given by the equation

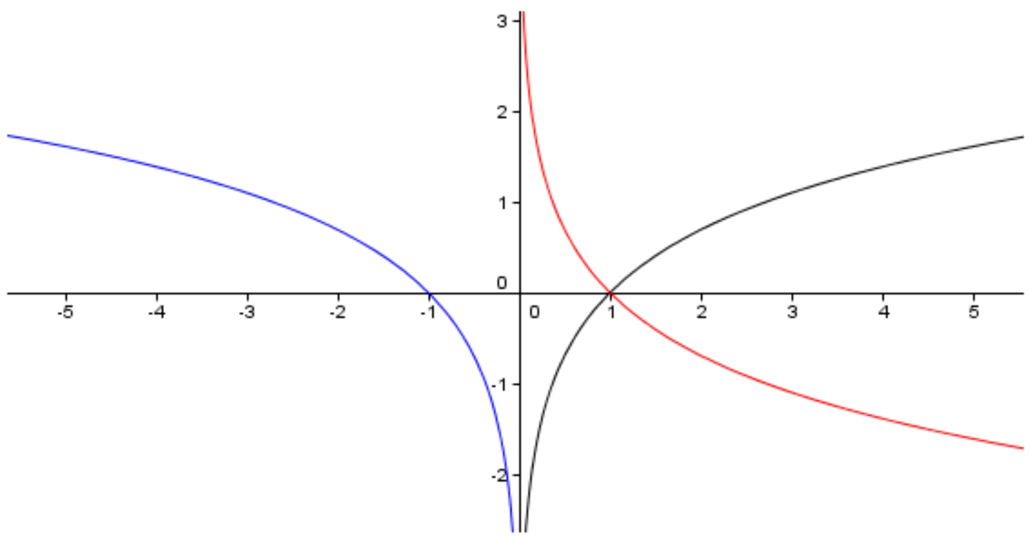
$y = -f(x)$

In this general equation, all y values are switched to their negative counterparts while the x values remain the same. The result is that the curve is flipped over the x-axis.

A horizontal reflection is given by the equation

$y = f(-x)$

Figure 2.50 Reflected Functions



The black line represents the original function, while the red and blue lines are vertical and horizontal reflections, respectively.

The third type of reflection is a reflection across the line $y = x$. This reflection has the effect of swapping the variables x and y, which is exactly what happens in the case of an inverse function. In [Figure 2.50](), the red and blue curves are the inverse of each other.

Source: https://www.boundless.com/algebra/graphs-functions-and-models/transformations/reflections/

Stretching and Shrinking

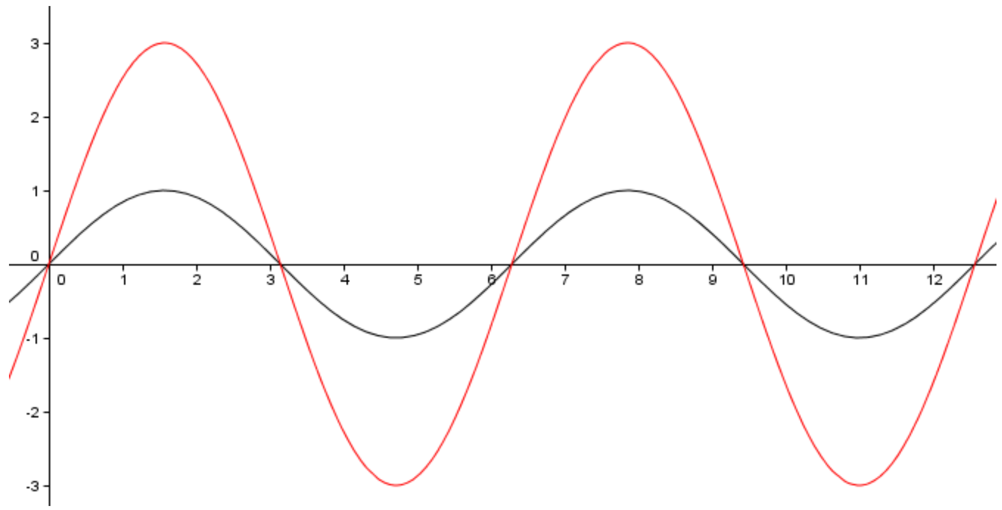
Stretching and shrinking refer to transformations that alter how compact a function looks in the x or y direction.

KEY POINTS

- When by either f(x) or x is multiplied by a number, functions can "stretch" or "shrink" vertically or horizontally, respectively, when graphed.
- In general, a vertical stretch is given by the equation y = pf(x). If p>1, the graph stretches upward and downward. If p<1, the graph shrinks.
- In general, a horizontal stretch is given by the equation y = f(x/q). If q>1, the graph shrinks horizontally, becoming more compact. If q<1, the graph stretches horizontally.

In algebra, equations can be stretched horizontally or vertically along an axis by multiplying either x or y by a number, respectively. By multiplying f(x) by a number greater than one, all the y values of an equation increase. This leads to a "stretched" appearance in the vertical direction. If f(x) is multiplied by a value less than one, all the y values of the equation decrease, leading to a "shrunken" appearance in the vertical direction. Alternatively, if only x is multiplied, the graph stretches or shrinks in the horizontal direction.

Figure 2.51 Sine Function Stretching Vertically



Here, the basic sine function is shown in black while the red function is stretched vertically by a factor of 3.

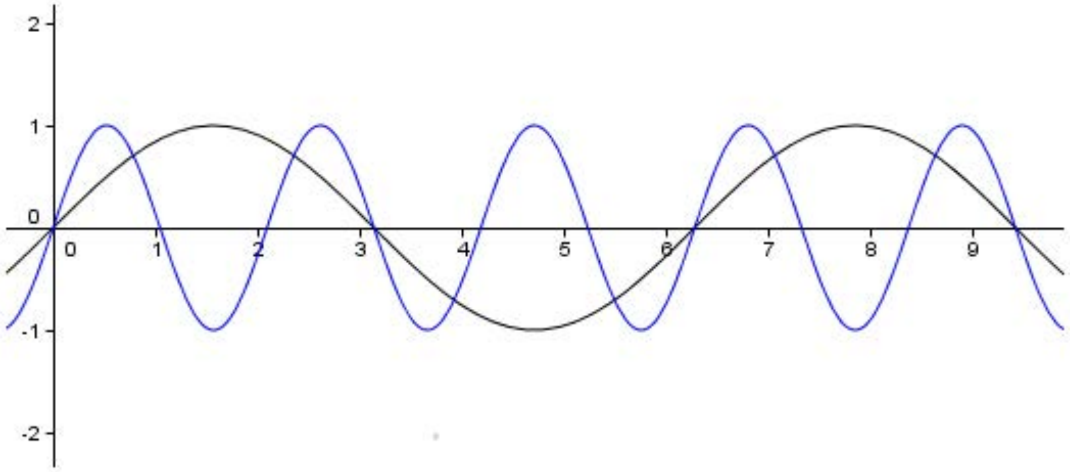
For examples, we will use the basic trignometric function f(x) = sin(x), which is black in the two graphs in *Figure 2.51*. Stretches can be a bit confusing with linear or **quadratic** functions, but they are much more straight forward with the sine function. The red function in *Figure 2.51* has been stretched (dilated) vertically by a factor of 3 and follows the equation:

$$y = 3f(x) = 3\sin x.$$

In general a vertical stretch is given by the equation:

$$y = pf(x)$$

Figure 2.52 Sine Function Stretching Horizontally



Here, the basic sine function is shown in black while the blue function is stretched horizontally by a factor of 3.

If p is larger than 1, the function gets "taller." If p is smaller than 1, the function gets "shorter."

The blue function in *Figure 2.52* has been been stretched horizontally by a factor of 3 and has the equation:

$y = f(3x) = \sin(3x)$.

In general, a horizontal stretch is given by the equation:

$y = f(x/q)$

In the example above, q = 1/3. When q is larger than 1, the function will get "longer" and when q is smaller than 1, the function will "squish."

Chapter 3

Functions, Equations, and Inequalities

https://www.boundless.com/algebra/functions-equations-and-inequalities/

Section 1

Linear Equations and Functions

https://www.boundless.com/algebra/functions-equations-and-inequalities/linear-equations-and-functions/

Linear Equations and their Applications

Linear equations are those with one or more variables of the first order.

KEY POINTS

- Linear equations can be expressed in the form: Ax+By+Cz+...=D.
- Linear equations can contain one or more variables; it's possible for such an equation to include an infinite number of variables.
- Linear equations can be used to solve for unknowns in any relationship in which all the variables are first order.

7 shelves? 2 * 10 + 7 * 5 = 55 Also to make it total from the 45 feet you also need 10 for a top and bottom.. So the actual answer should be 3 shelves. 2 * 10 for the sides 2 * 5 for the top and bottom and 3 * 5 for the shelves

A **linear equation** is an algebraic equation that is of the first order—that is, an equation in which each term is either a constant or the product of a constant and a variable raised to the first power.

Linear equations are commonly seen in two dimensions, but can be represented with three, four, or more variables. There is in fact a field of mathematics known as linear algebra, in which linear equations in up to an infinite number of variables are studied.

Linear equations can therefore be expressed in general (standard) form as:

$$ax + by + cz + \ldots = d$$

where a, b, c, and d are constants and x, y, and z are variables. Note that there can be infinitely more terms. This is known as general (or standard) form.

Applications of Linear Equations

Linear equations can be used to solve many problems, both everyday and technically specific.

Consider, for example, a situation in which one has 45 feet of wood to use for making a bookcase. If the height and width are to be 10 feet and 5 feet, respectively, how many shelves can be made between the top and bottom of the frame?

To solve this equation, we can use a linear relationship:

$$Nv + Mh = 45$$

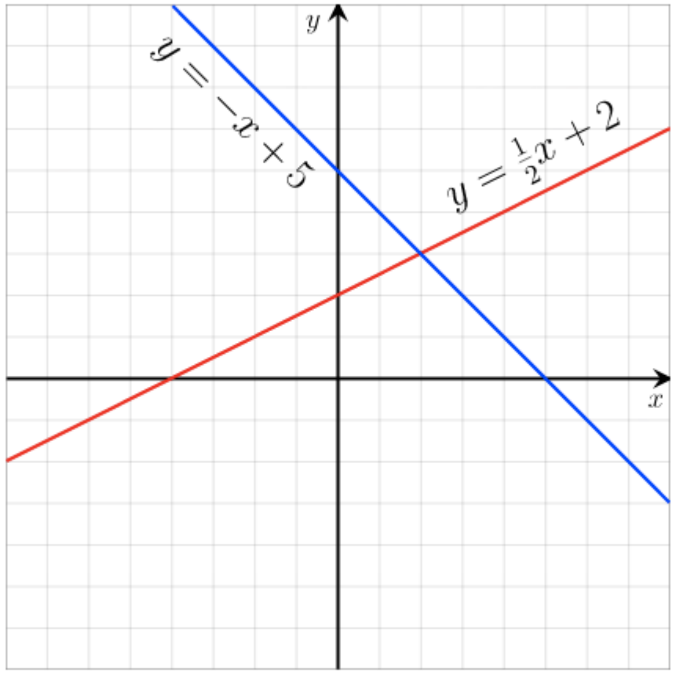


Figure 3.1 Linear Function Graph

These linear equations can represent the trajectories of two vehicles. If the drivers want to designate a meeting point, they can algebraically find the point of intersection of the two functions.

where v and h respectively represent the length in feet of vertical and horizontal sections of wood. N and M represent the number of vertical and horizontal pieces, respectively. Knowing that there will be only two vertical pieces, this formula can be simplified to:

$2 \cdot 10 + M \cdot 5 = 45$

Solving for M, we find that there is enough material for 5 shelves.

Similarly, we can use linear equations to solve for the original price of an item that is on sale. For example, consider an item that costs $24 when on a 40% discount. If the original price is x, we can write the following relationship:

$x - 0.4 \cdot x = 24$

Solving for x, we find that the original price was $40.

Using similar models we can solve equations pertaining to distance, speed, and time (Distance=Speed*Time); density (Density=Mass/Volume); and any other relationship in which all variables are first order.

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/linear-equations-and-functions/linear-equations-and-their-applications/

Zeros of Linear Functions

A zero, or x-intercept, is the point at which a linear function's value will equal zero.

KEY POINTS

- A zero is a point at which a function's value will be equal to zero. Its coordinates are (x, 0), where x is equal to the zero of the graph.
- Zeros can be observed graphically or solved for algebraically.
- A linear function can have zero, one, or infinitely many zeros. If the function is a horizontal line (slope=0), it will have no zero unless its equation is y=0, in which case it will have infinitely many. If the line is non-horizontal, it will have one zero.

A **linear function**, when graphed, is a straight line. It may or may not have an independent variable with a nonzero coefficient. If it does, that variable is of the first order; that is, the variable's exponent is equal to 1.

An x-intercept, or **zero**, is a property of many functions. This is a point at which the function crosses the x-axis. Thus, that point will have the value (x, 0), where x is the zero.

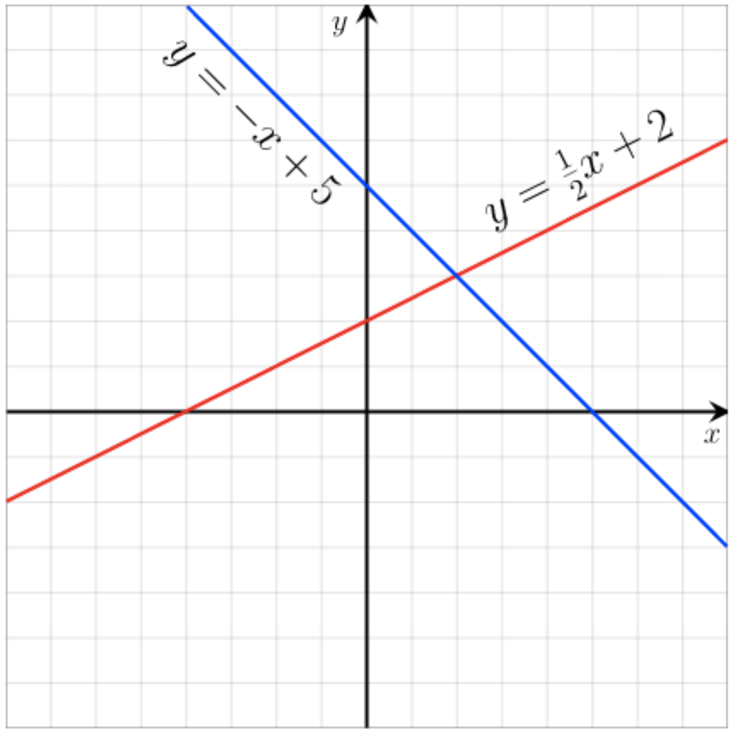


Figure 3.2 Linear Function Graph

Graphically, it can be observed that the equation y=-x +5 has a zero at x=5, and that y=1/2x+2 has a zero at x=-4

All lines will have one zero, provided they are functions of x to the first power, with the x-term having a nonzero coefficient. If the x-term has a zero coefficient, the line is horizontal and does not have an x-intercept unless the equation is y=0. If the equation is y=0, then all real values of x are valid x-intercepts.

Zeros can be observed graphically (*Figure 3.2*). Additionally, they can be computed algebraically.

Linear equations can be expressed in many forms, some of which are more explicit than others in revealing the x- and **y-intercept**s. The general (standard) form for a line is:

$$Ax + By = C$$

where x and y are variables and A, B, and C are constants.

In this form, the ratio of C/A is equal to the x-intercept (zero), so long as A is not equal to 0. If A is 0, the line is horizontal.

Lines can also be expressed in intercept form:

$$\frac{x}{a} + \frac{y}{b} = 1$$

In this case, the x-intercept (zero) is a and the y-intercept is b.

Another common form is slope-intercept form:

$$y = mx + b$$

where x and y are variables, m is the slope of the line and b is the y-intercept. To find the x-intercept, set y=0 (as it is at the point of the zero) and solve for x. Rearranging the formula gives:

$$x = \frac{-b}{m}$$

Thus, the x value at the zero will be equal to the quotient of the opposite of the y-intercept and the slope.

Formulas and Problem Solving

Linear equations can be used to solve many everyday and technically specific problems.

KEY POINTS

- A linear equation can be used to solve any problem that includes constants and variable(s) of first order.
- A linear equation can be solved for any one variable provided that the values of all other variables are known.
- Linear equations can be used to calculate tip, cost of goods, velocity, simple interest, and many more variables.

Linear equations can be used to solve many practical and technical problems. Such an equation may include many variables so long as all are of the first order, and the value of any one variable can be calculated if the values of all the other variables are known.

For example, one can use a linear equation to determine the amount of interest accrued on a home equity line of credit after a given amount of time. Consider the hypothetical situation in which you need money to make home improvements and can open a $20,000 credit line at an interest rate of 2.5% per year. You plan to pay off the debt in its entirety within 15 months. To find out how much it will cost you can use following formula:

$$I = P \cdot r \cdot T$$

Where I is interest, p is the principal amount loaned ($20,000), r is the interest rate (2%, or 0.02) per year, and T is the number of years elapsed (15 months will be 1.25 years).

Plugging the known values into the above formula, we can determine that you will pay $500 in interest.

There are many other common formulas that can be used for everyday computations. Some have more variables than others, but none has a variable of order higher than one. Let's take a few examples of other linear equations, namely velocity, gratuity (tip), and cost of purchased goods:

$$V = \frac{d}{T}$$

The formula relating velocity (V), distance (d), and time (T).

$G = c \cdot r$ (Figure 3.3)

The formula relating gratuity (G), cost (c), and desired percent gratuity (r, expressed as a decimal).

$$A \cdot x + B \cdot y + C \cdot z + \ldots = T$$

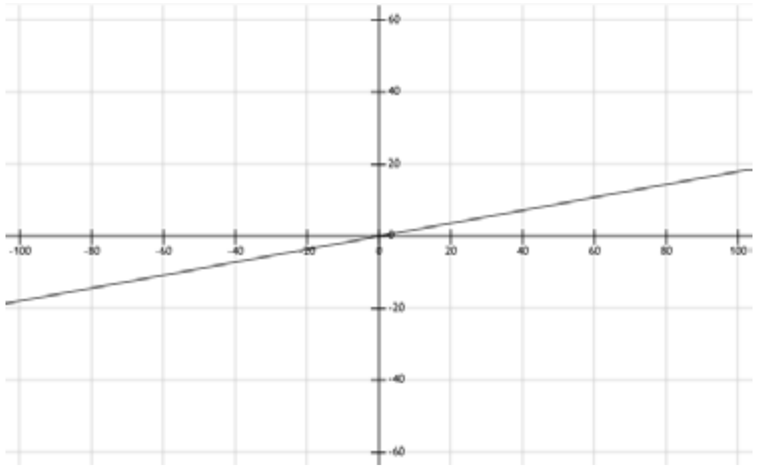


Figure 3.3 Gratuity as a function of bill price

In the above plot, the dependent variable (y) represents gratuity (tip) as a function of cost of the bill (x) before gratuity. The tip rate is 18%.

Where A, B and C represent the quantities of goods that cost x, y and z, respectively. This could be expanded or contracted relative to the number of different items purchased. T represents the total cost of goods purchased.

Section 2

Complex Numbers

https://www.boundless.com/algebra/functions-equations-and-inequalities/complex-numbers/

The Complex-Number System

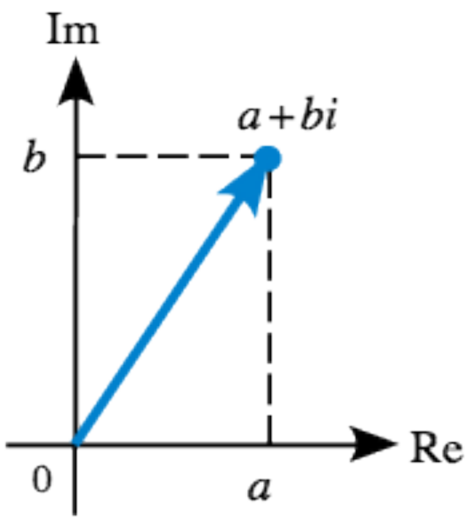
A complex number has the form a + bi, where a and b are real numbers and i is the imaginary unit.

KEY POINTS

- A complex number is a number that can be expressed in the form a + bi, where a and b are real numbers and i is the imaginary unit.
- The real number a of the complex number z = a + bi is called the real part of z, and the real number b is often called the imaginary part.
- The real part is denoted by Re(z) or $\Re(z)$, and the imaginary part b is denoted by Im(z) or $\Im(z)$.

A **complex** number is a number that can be put in the form a + bi, where a and b are **real numbers** and i is called the imaginary unit, where i2 = −1. In this expression, a is called the real part and b the imaginary part of the complex number. Complex numbers extend the idea of the one-dimensional number line to the two-dimensional complex plane by using the horizontal axis for the real part and the vertical axis for the imaginary part. The complex number a + bi can be identified with the point (a, b) as shown in the figure below. A complex number whose real part is zero is said to be purely imaginary, whereas a complex number whose imaginary part is zero is a real number. In this way, complex numbers contain the ordinary real numbers while extending them in order to solve problems that cannot be solved with only real numbers (Figure 3.4).

Figure 3.4 Complex number illustration



A complex number can be visually represented as a pair of numbers (a,b) forming a vector on the complex plane. "Re" is the real axis, "Im" is the imaginary axis, and "i" is the imaginary unit, satisfying i^2 = −1.

Complex numbers are used in many scientific fields, including engineering, electromagnetism, quantum physics, and applied mathematics, such as chaos theory. Italian mathematician Gerolamo Cardano is the first known to have introduced complex numbers. He called them "fictitious" during his attempts to find solutions to cubic equations in the 16th century.

Complex numbers allow for solutions to certain equations that have no real solution. For example, the equation

$$(x+1)^2 = -9$$

has no real solution, since the square of a real number is either 0 or positive. Complex numbers provide a solution to this problem. The idea is to extend the real numbers with the imaginary unit i where $i2 = -1$, so that solutions to equations like the preceding one can be found. In this case, the solutions are $-1 \pm 3i$. In fact, not only quadratic equations, but all polynomial equations in a single variable can be solved using complex numbers.

Another is example is the complex number $-3.5 + 2i$. It is common to write a for $a + 0i$ and bi for $0 + bi$. Moreover, when the imaginary part is negative, it is common to write $a - bi$ with $b > 0$ instead of $a + (-b)i$, for example $3 - 4i$ instead of $3 + (-4)i$.

The real number a of the complex number $z = a + bi$ is called the real part of z, and the real number b is often called the imaginary part. By this convention, the imaginary part is a real number – not including the imaginary unit: hence b, not bi, is the imaginary part. The real part is denoted by Re(z) or $\Re(z)$, and the imaginary part b is denoted by Im(z) or $\Im(z)$.

Some authors write $a + ib$ instead of $a + bi$ (scalar multiplication between b and i is commutative).

A real number a can usually be regarded as a complex number with an imaginary part of zero, that is to say, $a + 0i$. A pure **imaginary number** is a complex number whose real part is zero, that is to say, of the form $0 + bi$.

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/complex-numbers/the-complex-number-system/

Addition and Subtraction, and Multiplication

Complex numbers are added by adding the real and imaginary parts; multiplication follows the rule $i^2 = -1$.

KEY POINTS

- Complex numbers are added by adding the real and imaginary parts of the summands. That is to say: (a + bi) + (c + di) = (a + c) + (b + d)i.
- Similarly, subtraction is defined by (a + bi) - (c + di) = (a - c) + (b - d)i.
- The multiplication of two complex numbers is defined by the following formula: (a + bi)(c + di) = (ac - bd) + (bc + ad)i.

Addition and Subtraction

Complex numbers are added by adding the real and imaginary parts of the summands. That is to say:

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

Similarly, subtraction is defined by:

$$(a + bi) - (c + di) = (a - c) + (b - d)i$$

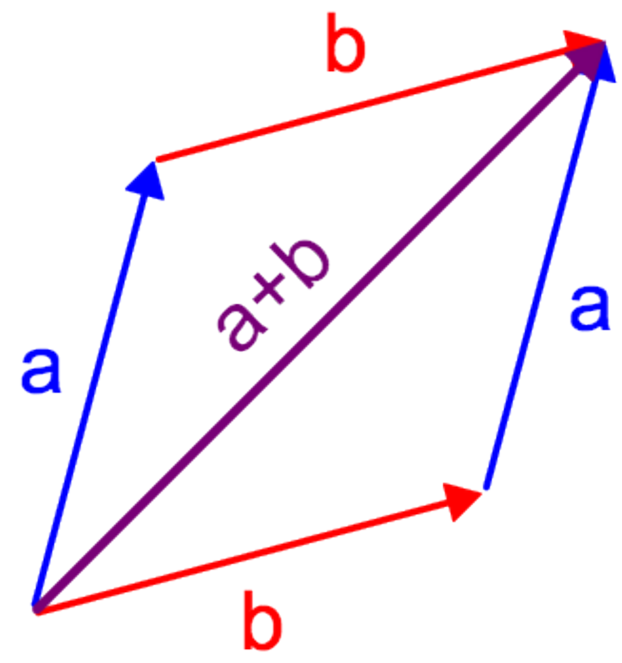


Figure 3.5 Addition of complex numbers

Addition of two complex numbers can be done geometrically by constructing a parallelogram.

Using the visualization of complex numbers in the complex plane, the addition has the following geometric interpretation: the sum of two complex numbers A and B, interpreted as points of the complex plane, is the point X obtained by building a **parallelogram**, three of whose vertices are O, A, and B (as shown in *Figure 3.5*).

Multiplication

The multiplication of two complex numbers is defined by the following formula:

$$(a + bi)(c + di) = (ac - bd) + (bc + ad)i$$

In particular, the square of the **imaginary unit** is -1:

$i^2 = i \cdot i = -1$

The preceding definition of multiplication of general complex numbers follows naturally from this fundamental property of the imaginary unit. Indeed, if i is treated as a number so that di means d time i, the above multiplication rule is identical to the usual rule for multiplying the sum of two terms.

$(a + bi)(c + di) = ac + bci + adi + bidi$ (by the distributive law)

$= ac + bidi + bci + adi$ (by the commutative law of addition)

$= ac + bdi^2 + (bc + ad)i$ (by the commutative law of multiplication)

$= (ac - bd) + (bc + ad)i$ (by the fundamental property of the imaginary unit)

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/complex-numbers/addition-and-subtraction-and-multiplication/

Complex Conjugates and Division

The complex conjugate of x + yi is x - yi, and the division of two complex numbers can be defined using the complex conjugate.

KEY POINTS

- The complex conjugate of the complex number z = x + yi is defined to be x - yi. It is denoted z*.
- The division of two complex numbers is defined in terms of complex multiplication and real division. Where at least one of c and d is non-zero: (a + bi)/(c + di) = (ac + bd)/(c2 + d2) + [(bc - ad)/(c2 + d2)]i.
- Division can be defined in this way because of the following observation: (a + bi)/(c + di) = [(a + bi)(c - di)]/[(c + di)(c - di)] = (ac + bd)/(c2 + d2) + [(bc - ad)/(c2 + d2)]i. As shown earlier, c - di is the complex conjugate of the denominator c + di.

Conjugation

The **complex conjugate** of the complex number z = x + yi is defined as x - yi. It is denoted z*. Geometrically, z* is the "reflection" of z about the real axis (as shown in the figure below).

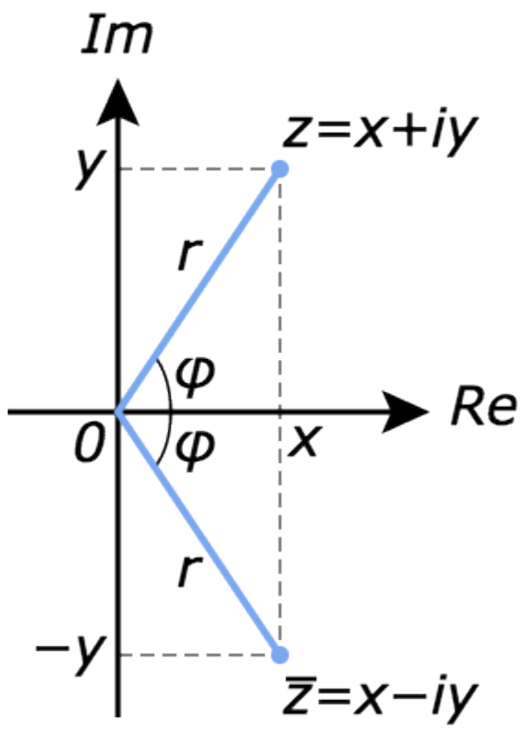


Figure 3.6 Complex Conjugate

Geometric representation of z and its conjugate in the complex plane.

Specifically, conjugating twice gives the original complex number: z** = z (*Figure 3.6*).

The real and **imaginary** parts of a complex number can be extracted using the conjugate, respectively:

$$Re(z) = (1/2)(z + z^*)$$

$$Im(z) = (1/2i)(z - z^*)$$

Moreover, a complex number is real if and only if it equals its conjugate.

Conjugation distributes over the standard arithmetic operations:

$$(z + w)^* = z^* + w^*$$

$$(zw)^* = z^* w^*$$

$$(z/w)^* = z^*/w^*$$

The reciprocal of a nonzero complex number $z = x + yi$ is given by

$$1/z = z^*/zz^* = z^*/(x^2 + y^2)$$

Division

The division of two complex numbers is defined in terms of complex multiplication (described above) and real division. Where at least one of c and d is non-zero:

$$\frac{(a+bi)}{(c+di)} = \frac{(ac+bd)}{(c^2+d^2)} + \frac{(bc-ad)}{(c^2+d^2)}i$$

Division can be defined in this way because of the following observation:

$$\frac{(a+bi)}{(c+di)} = \frac{(a+bi)(c-di)}{(c+di)(c-di)} = \frac{(ac+bd)}{(c^2+d^2)} + \frac{(bc-ad)}{(c^2+d^2)}i$$

As shown earlier, c - di is the complex conjugate of the **denominator** c + di. Neither the real part c nor the imaginary part d of the denominator can be equal to zero for division to be defined.

Quadratic Equations, Functions, and Applications

https://www.boundless.com/algebra/functions-equations-and-inequalities/quadratic-equations-functions-and-applications/

Quadratic Equations and Quadratic Functions

Equations that are quadratic are of the second order, and have the form $f(x) = ax^2 + bx + c$.

KEY POINTS

- A quadratic function is of the form: $f(x) = ax^2 + bx + c$ where a is a nonzero constant, b and c are constants of any value, and x is a variable. A quadratic equation is a quadratic function set equal to 0.
- The shape of a quadratic function is known as a parabola.
- The solution(s) to a quadratic equation is/are known as its zero(s), or root(s).

A **quadratic function** is of the form:

$$f(x) = ax^2 + bx + c$$

where a, b, and c are constants and x is a variable. The constants b and c can take any value, but a cannot be equal to 0.

The single defining feature of **quadratic** functions is that they are of the second order (degree). Therefore, in all quadratic functions the highest exponent of x in a nonzero term is equal to 2.

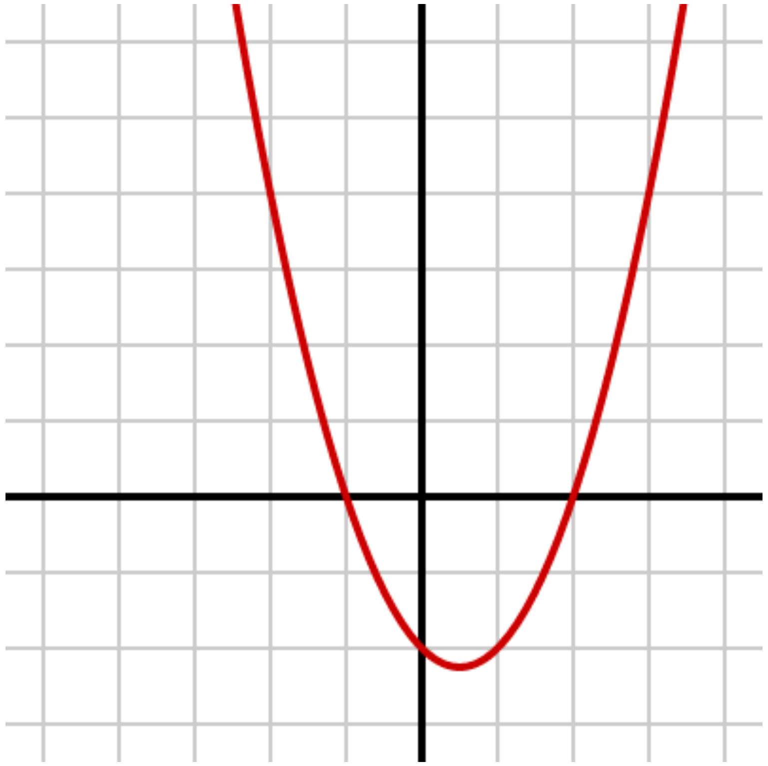


Figure 3.7
Polynomial
The function f(x)=x^2-x-2 is quadratic; it has a parabolic shape and is of the second order.

The shape of a quadratic function is called a **parabola**. If a is positive, the shape resembles a U; if a is negative, the U is flipped upside-down.

Quadratic functions can be expressed in many different forms. The form written above is standard form. Additionally:

$$f(x) = a(x - x_1)(x - x_2)$$

is known as factored form, where x_1 and x_2 are the zeros, or roots, of the equation. These are x values at which the function crosses the y-axis (and thus y equals 0). Also:

$$f(x) = a(x - h)^2 + k$$

is known as the **vertex** form, where h and k are respectively the coordinates of the vertex, the point at which the function reaches either its maximum (if a is negative) or minimum (if a is positive).

A quadratic equation is a specific case of a quadratic function, with the function set equal to 0:

$$ax^2 + bx + c = 0$$

As in a quadratic function, x is a variable, a, b, and c are constants, and a cannot equal 0.

With a, b, and c known, a quadratic equation can be solved for x, such solutions are known as zeros. There are several ways of finding x, which depending on the equation can have 0, 1, or 2 values, but these methods will be discussed later.

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/quadratic-equations-functions-and-applications/quadratic-equations-and-quadratic-functions/

Completing the Square

Completing the square is a common method for solving quadratic equations: $0 = a(x-h)^2 + k$.

KEY POINTS

- In the form, $a(x-h)^2 + k$, h and k are constants and should be found.
- Once a quadratic polynomial is in the form: $0 = a(x-h)^2 + k$, one can solve for two values of x (using the positive and negative square roots).
- When a parabola is plotted, h is the x-coordinate of the axis of symmetry, and k is the minimum value (or maximum value, if a < 0) of the quadratic function.

Along with factoring and using the quadratic formula, completing the square is a common method for solving quadratic equations.

Consider a generic quadratic equation:

$$0 = ax^2 + bx + c$$

"Completing the square," one can convert that equation to the form:

$$0 = a(x-h)^2 + k$$

where one must find constants h and k.

This method is meant to be used when the quadratic function is not a perfect square; the value of k is meant to adjust the function to compensate for the difference between the expanded a(x-h)2 and the quadratic function ax2+bx+c.

There is a simple formula in elementary algebra for computing the square of a binomial:

$$(x+p)^2 = x^2 + 2px + p^2$$

For example:

$$(x+3)^2 = x^2 + 6x + 9 \qquad (p = 3)$$

$$(x-5)^2 = x^2 - 10x + 25 \qquad (p = -5)$$

In any perfect square, the number p is always half the coefficient of x, and the constant term is equal to p2.

Consider the following quadratic polynomial:

$$x^2 + 10x + 28$$

This quadratic is not a perfect square, since 28 is not the square of 5:

$$(x+5)^2 = x^2 + 10x + 25$$

However, it is possible to write the original quadratic as the sum of this square and a constant:

$$x^2 + 10x + 28 = (x+5)^2 + 3$$

Thus, -5 is equal to h and 3 is equal to k.

Knowing that:

$$(x+5)^2 + 3 = x^2 + 10x + 28 = 0$$

We can solve for x:

$$(x+5)^2 = 3$$

$x + 5 = -\sqrt{3}$ and $x + 5 = \sqrt{3}$

Thus:

$x = -\sqrt{3} - 5$ and $x = \sqrt{3} - 5$

Let's go through the steps of completing the square another quadratic equation:

$$3x^2 + 12x + 27 = 3(x^2 + 4x + 9)$$

Factor out the coefficient a, and then complete the square for the resulting monic polynomial:

$$= 3[(x+2)^2 + 5]$$

$= 3(x+2)^2 + 15$

Which brings us back to the form:

$a(x-h)^2 + k$

where h=-2 and k=15.

Graphical Representation

The graph of any quadratic function is a **parabola** in the xy-plane. Given a quadratic polynomial of the form:

$a(x-h)^2 + k$

the numbers h and k may be interpreted as the Cartesian coordinates of the **vertex** of the parabola. That is, h is the x-coordinate of the axis of symmetry, and k is the minimum value (or maximum value, if a < 0) of the quadratic function.

In other words, the graph of the function $f(x) = x^2$ is a parabola whose vertex is at the origin (0, 0). Therefore, the graph of the function $f(x-h) = (x-h)^2$ is a parabola shifted to the right by h whose vertex is at (h, 0), as shown in *Figure 3.8*. In contrast, the graph of the function $f(x) + k = x^2 + k$ is a parabola shifted upward by k whose vertex is at (0, k), as shown in *Figure 3.9*. Combining both horizontal and vertical shifts yields $f(x-h) + k = (x-h)^2 + k$ is a parabola shifted to the right by h and upward by k whose vertex is at (h, k), as shown in *Figure 3.10*.

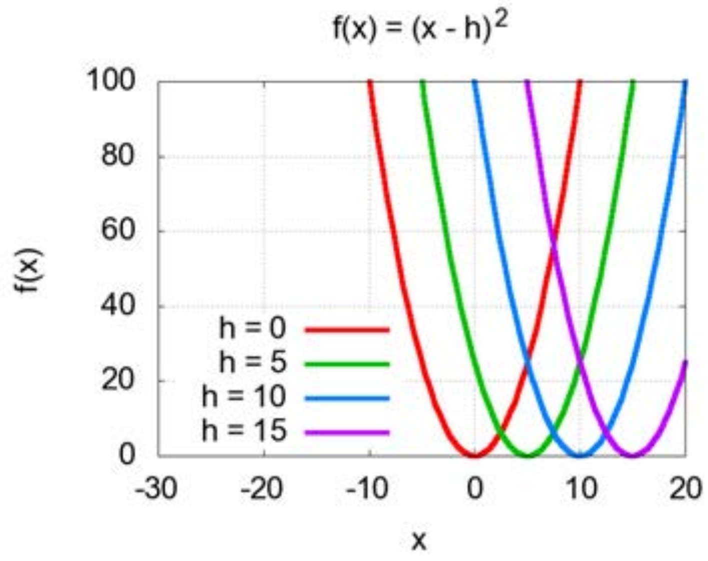


Figure 3.8 H values in Quadratic Functions

Graphs of quadratic functions shifted to the right by h = 0, 5, 10, and 15.

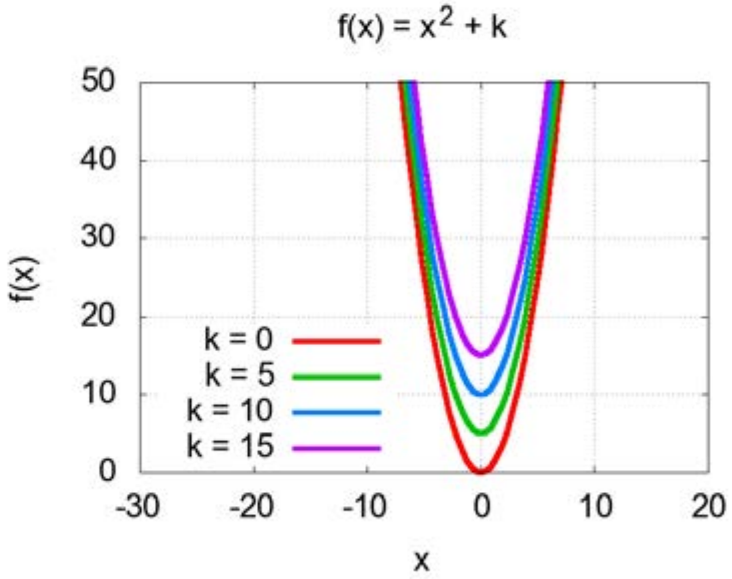


Figure 3.9 K values in Quadratic Functions

Graphs of quadratic functions shifted upward by k = 0, 5, 10, and 15.

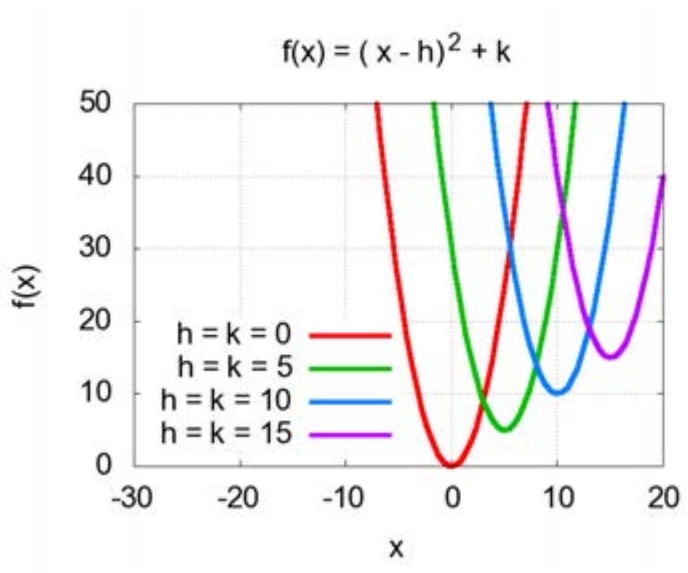


Figure 3.10 Quadratic functions at varying h and k values

Graphs of quadratic functions shifted upward and to the right by 0, 5, 10, and 15.

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/quadratic-equations-functions-and-applications/completing-the-square/

The Quadratic Formula

The zeros of a quadratic equation can be found not only through factoring, but by solving what is known as the quadratic formula.

KEY POINTS

- The quadratic formula is: $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ where a, and b are the coefficients of the x^2 and x terms, respectively, in a quadratic equation, and c is the value of the equation's constant.
- If the discriminant (b^2-4ac) is equal to zero, both values of x obtained will be the same, real number.
- If the discriminant is positive, both roots will be distinct and real.
- If the discriminant is negative, there will be two imaginary roots with i coefficients.

Recall that a quadratic equation essentially reveals one or two points of a quadratic function. Whereas a quadratic function will have the form:

$$f(x) = ax^2 + bx + c$$

with infinite solutions possible for f(x) and x, a quadratic equation will have the form:

$$0 = ax^2 + bx + c$$

In a quadratic equation, there is only one unknown value (x), not two (f(x) and x). As such, one can solve for x, which happens to be the value of the x-intercept(s), also known as roots or **zero**s. If f(x)=0 at the vertex of the **parabola**, there will be one value of x that fits; otherwise, there will be two.

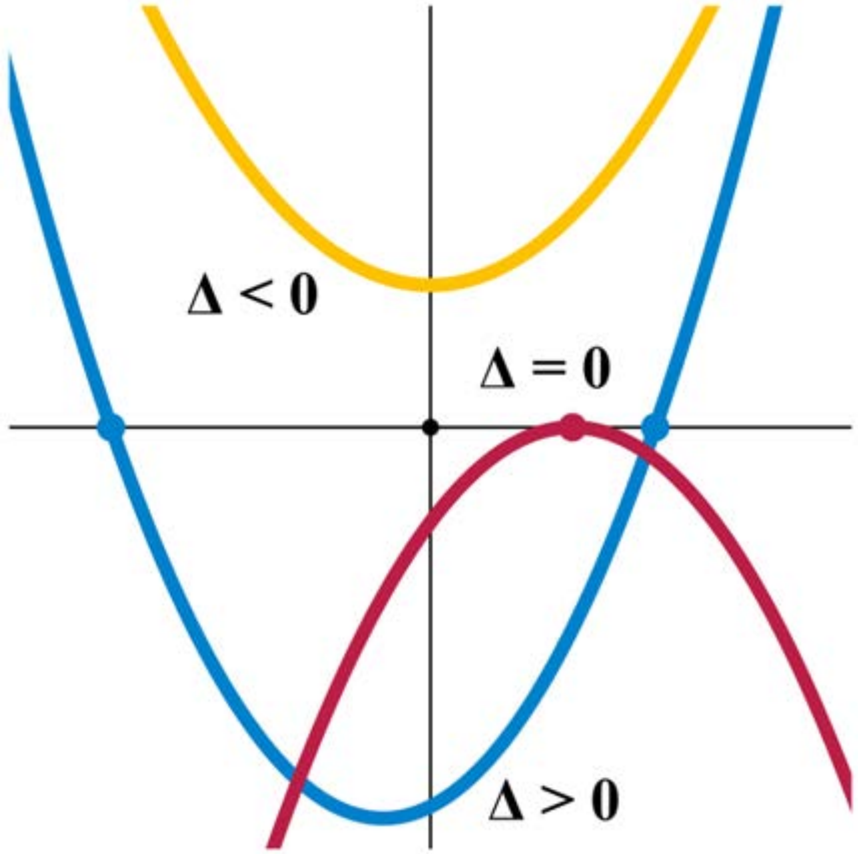
Finding X

X can be found not only by factoring, but also with use of the quadratic formula. That is:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

where the symbol ± indicates there will be two solutions, one that involves adding the square root of the discriminant (b2-4ac), and the other found by subtracting said square root. The resulting x values (zeros) may or may not be distinct, and may or may not be real.

Figure 3.11 Parabolas of differing discriminant values



Note that the function with discriminant (Δ) value greater than 0 crosses the x-axis twice, explaining the two distinct values of zeros obtained from the quadratic formula. When the discriminant value is 0, there is only one point at which the function touches the x-axis; hence there being only one calculable root. And for discriminants less than 0, the functions never touch the x-axis at all, resulting in no real roots, only imaginary ones.

Determining values

If the discriminant is equal to zero, both values of x obtained will be the same, real number. If the discriminant is positive, both roots

will be distinct and real. If the discriminant is negative, there will be two imaginary roots, defined by:

$$\frac{-b}{2a} + i\frac{\sqrt{-(b^2 - 4ac)}}{2a}$$

and

$$\frac{-b}{2a} - i\frac{\sqrt{-(b^2 - 4ac)}}{2a}$$

All the possibilities concerning number of solutions to a quadratic equation can be explained by examining the graphs of parabolas with different discriminant values (*Figure 3.11*).

The Discriminant

The discriminant of a polynomial is a function of its coefficients that reveals information about the polynomial's zeros, or roots.

KEY POINTS

- $\Delta = b^2 - 4ac$ is the formula for a quadratic expression's discriminant, where Δ is the discriminant, and a, b and c are coefficients from the parent function.
- If Δ is greater than 0, the polynomial has two real, distinct zeros. If Δ is equal to 0, the polynomial has only one real zero. And if Δ is less than 0, the polynomial has no real zeros, only two distinct complex zeros.
- A zero is the x value whereat the function crosses the x-axis. That is, it is the x-coordinate at which the function's value equals 0.

The **discriminant** of a polynomial is a function of its coefficients that reveals information about the polynomial's **zero**s, or roots. A zero is the x value whereat the function crosses the x-axis. That is, it is the x-coordinate at which the function's value equals 0.

The concept of a discriminant can be applied to many different orders of polynomial; depending on degree, the discriminant can be calculated with differing equations. In this case, we will discuss its

application to those polynomials of the second order (**quadratics**). Recall the form of a quadratic polynomial:

$$ax^2 + bx + c$$

where a, b and c are constants (a must be nonzero), and x is a variable.

The discriminant (Δ) of the above polynomial can be calculated based on the equation:

$$\Delta = b^2 - 4ac$$

If Δ is greater than 0, then the polynomial has two real, distinct zeros. If Δ is equal to 0, then the polynomial has only one real zero. And if Δ is less than 0, then the polynomial has no real zeros; it has only two distinct complex zeros.

Example

Consider, for example, the quadratic function:

$$f(x) = x^2 - x - 2$$

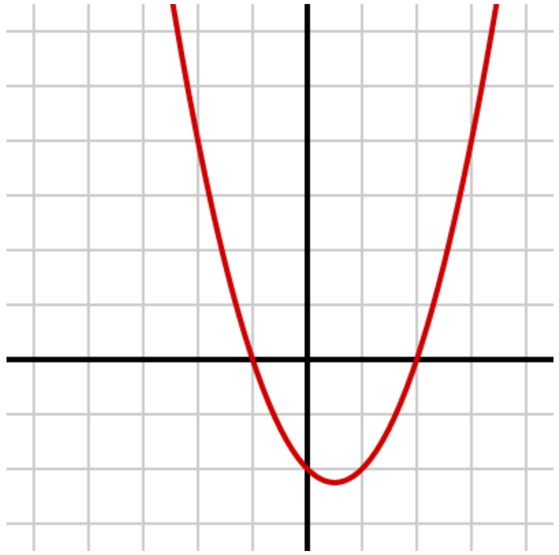
Using 1 as the value of a, -1 as the value of b, and -2 as the value of c, the discriminant of this function can be determined as follows:

$$\Delta = (-1)^2 - 4 \cdot 1 \cdot (-2)$$

$$\Delta = 9$$

Because Δ is greater than 0, the function has two distinct, real zeros. Checking graphically, we can confirm this is true; the zeros of the function can be found at x=-1 and x=2 (*Figure 3.12*).

Figure 3.12 Polynomial



The discriminant of f(x)=x^2-x-2 is 9. Because the value is greater than 0, the function has two distinct, real zeros. The graph of f(x) shows that it clearly has two roots: the function crosses the x-axis at x=-1 and x=2.

Reducing Equations to a Quadratic

Quartic equations with no terms of odd-degree can be reduced to quadratics; the square roots of their solutions solve the parent equation.

KEY POINTS

- A biquadratic equation (quartic equation with no terms of odd-degree) has the form $0 = ap^2 + bp + c$ (where p=x2).
- The values of p can be found by graphing, factoring, completing the square, or using the quadratic formula. Their square roots (positive and negative) are the values of x that satisfy the original equation.
- Higher-order equations can be solved by a similar process that involves reducing their exponents. The requirement is that there are two terms of x such that the ratio of the highest exponent of x to the lower is 2:1.

Quartic (fourth-degree) equations can be very difficult to solve. In some special situations, however, they can be more manageable.

If, for example, a quartic equation is biquadratic—that is, it includes no terms of an odd-degree— there is a quick way to find the zeroes of a quartic function by reducing it into a quadratic form.

Consider a quartic function with no odd-degree terms, which therefore has the form:

$$0 = ax^4 + bx^2 + c$$

Substituting p in place of x^2 (and thus p^2 in place of x^4), this can be reduced to an equation of lower degree:

$$0 = ap^2 + bp + c$$

This quadratic equation can be solved for p by any of a number of methods (by graphing, factoring, completing the square, or by using the quadratic formula).

Once values of p are found, each positive value thereof can be used to find two values of x as such:

$$x = \sqrt{p}$$

As with every square root, that of p will have two (positive and negative) values.

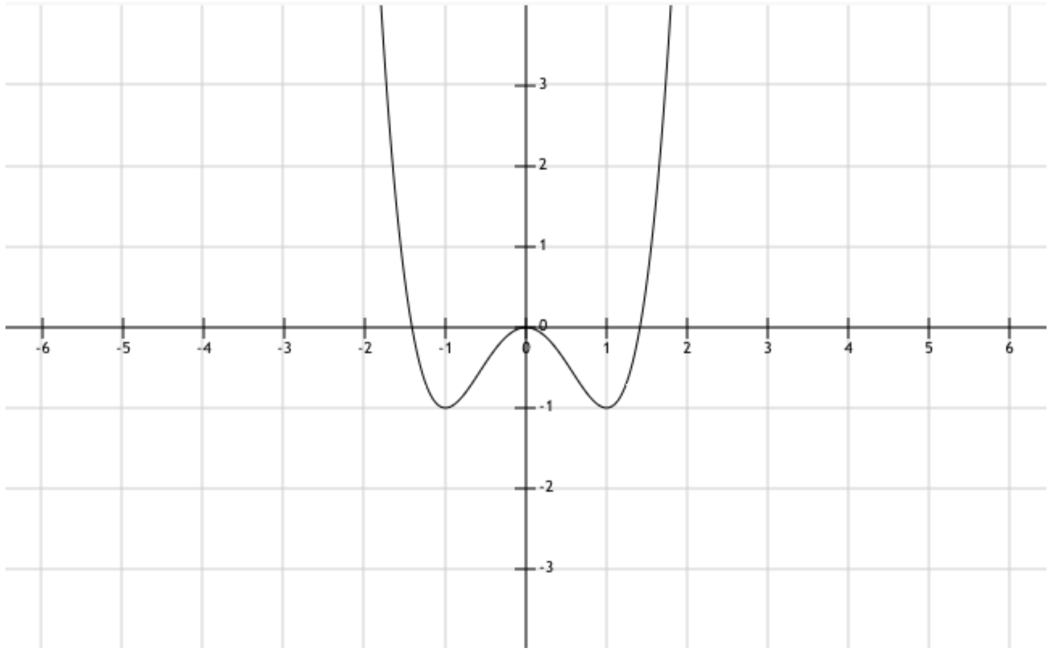
Example

Consider, for example, the quartic equation:

$$0 = x^4 - 12x^2 + 20$$

We can substitute p in place of x^2:

Figure 3.13 Graph of f(x)=x^4-2x^2



This equation is biquadratic and has three distinct zeroes that can be found by reducing the terms to quadratic form and finding their square roots.

$$0 = p^2 - 12p + 20$$

And solve for p using the quadratic formula:

$$p = \frac{12 \pm \sqrt{(-12)^2 - 4 \cdot 1 \cdot 20}}{2 \cdot 1}$$

$p = 2$ and $p = 10$

Knowing that p=x2, we can use each value of p to solve for two values of x:

$x = \pm\sqrt{2}$ and $x = \pm\sqrt{10}$

A similar procedure can be used to solve higher-order equations. The requirement is that there are two terms of x such that the ratio of the highest exponent of x to the lower is 2:1.

Applications and Problem Solving

Quadratic relationships between variables are commonly found in physical sciences, engineering, and elsewhere.

KEY POINTS

- The Pythagorean Theorem, $c^2 = a^2 + b^2$ relates the length of the hypotenuse (c) of a right triangle to the lengths of its legs (a and b).
- Problems involving gravity and projectile motion are typically dependent upon a second-order variable, usually time or initial velocity depending on the relationship.
- Coulomb's Law, which relates electrostatic force, charge amount and distance between two charged particles, has a second-order dependence on the separation of the particles. Solving for either charge results in a quadratic function.

Quadratic relationships between variables are commonly found in physical sciences, engineering, and elsewhere. Perhaps the most universally used example of quadratic relationships in problem solving concerns right triangles.

The Pythagorean Theorem

The Pythagorean Theorem is used to relate the three sides of right triangles. It states:

$$c^2 = a^2 + b^2$$

According thereto, the square of the length of the hypotenuse (c) is equal to the sum of the squares of the two legs (a and b) of the triangle. This has been proven in many ways, among the most famous of which was devised by Euclid (*Figure 3.14*).

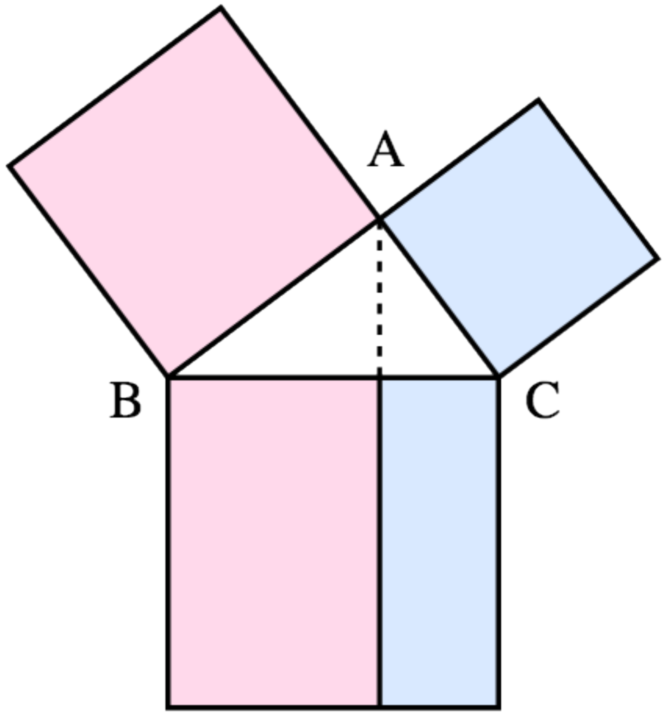


Figure 3.14
Euclid's Proof of the Pythagorean Theorem
Euclid used this diagram to explain how the sum of the squares of the triangle's smaller sides (pink and blue) sum to equal the area of the square of the hypotenuse.

Gravity and Projectile Motion

Most all equations involving gravity include a second-order relationship. Consider, for example, the equation relating gravitational force (F) between two objects to the masses of each object (m_1 and m_2) and the distance between them (r):

$$F = G\frac{m_1 m_2}{r^2}$$

The shape of this function is not a parabola, but becomes such a shape if rearranged to solve for m_1 or m_2.

The maximum height of a projectile launched directly upwards can also be calculated from a quadratic relationship. The formula relates height (h) to initial velocity (v_0) and gravitational acceleration (g):

$$h = \frac{v_0^2}{2g}$$

The same maximum height of a projectile launched directly upwards can be found using the time at the projectile's peak (t_h):

$$h = v_0 t_h \frac{1}{2} g t_h^2$$

Substituting any time (t) in place of t_h leaves the equation for height as a function of time.

Electrostatic Force

The form of the equation relating electrostatic force (F) between two particles, the particles' respective charges (q_1 and q_2), and the distance between them (r) is very similar to the aforementioned formula for gravitational force:

$$F = \frac{q_1 q_2}{4\pi\epsilon_0 r^2}$$

This is known as Coulomb's Law. Solving for either charge results in a quadratic equation where the charge is dependent on r^2.

Graphs of Quadratic Functions

Quadratic Functions of the Form: $f(x) = a(x - h)^2 + k$

Quadratic Functions of the Type $f(x) = ax^2 + bx + c$ Where a is not Equal to 0

Applications

https://www.boundless.com/algebra/functions-equations-and-inequalities/graphs-of-quadratic-functions/

Quadratic Functions of the Form: $f(x) = a(x-h)^2 + k$

The graph of a quadratic function is a parabola whose axis of symmetry is parallel to the y-axis.

KEY POINTS

- If the quadratic function is set equal to zero, then the result is a quadratic equation.
- The solutions to the equation are called the roots of the equation.
- The coefficient a controls the speed of increase (or decrease) of the quadratic function from the vertex.
- The coefficients b and a together control the axis of symmetry of the parabola.
- The coefficient b alone is the declivity of the parabola as y-axis intercepts.
- The coefficient c controls the height of the parabola, or more specifically, it is the point where the parabola intercepts the y-axis.

A **quadratic** function can be expressed in three formats:

$f(x) = ax^2 + bx + x$ is called the standard form.

$f(x) = a(x-x_1)(x-x_2)$ is called the factored form, where x_1 and x_2 are the roots of the quadratic equation (used in logistic map).

$f(x) = (x-h)^2 + k$ is called the **vertex** form, where h and k are the x and y coordinates of the vertex, respectively.

To convert the standard form to factored form, one needs only the quadratic formula to determine the two roots x_1 and x_2. To convert the standard form to vertex form, one needs a process called completing the square. To convert the factored form (or vertex form) to standard form, one needs to multiply, expand, and/or distribute the factors.

Looking at a general case in which b=0. *Figure 3.15* shows the dependency on the quadratic form.

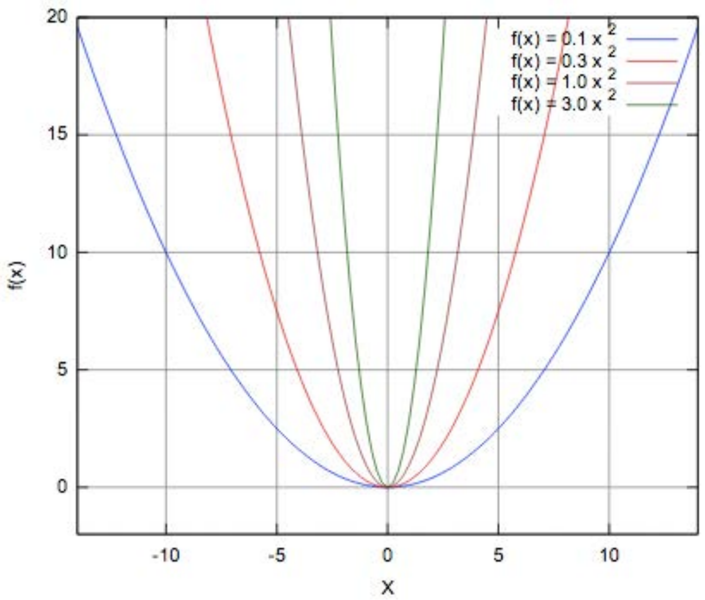


Figure 3.15 **Standard Polynomial**

This is a standard polynomial, with f(x) plotted against x, for f(x)=a2 for various values of a.

In elementary algebra, completing the square is a technique for converting a quadratic polynomial of the form

$$f(x) = ax^2 + bx + x$$

to the form

$$a(\ldots\ldots\ldots)^2 + constant$$

In this context, "**constant**" means not depending on x. The expression inside the parenthesis is of the form (x – constant). Thus one converts $ax^2 + bx + c$ to

$$f(x) = (x - h)^2 + k$$

and one must find h and k.

In analytic geometry, the minimum value graph of any quadratic function is a parabola in the xy-plane. Given a quadratic polynomial of the form

$$(x - h)^2 + k \quad or \quad a(x - h)^2 + k$$

the numbers h and k may be interpreted as the Cartesian coordinates of the vertex of the parabola. That is, h is the x-coordinate of the axis of symmetry, and k is the minimum value (or maximum value, if a < 0) of the quadratic function.

In other words, the graph of the function $f(x) = x^2$ is a parabola whose vertex is at the origin (0, 0). Therefore, the graph of the function $f(x - h) = (x - h)^2$ is a parabola shifted to the right by h and whose vertex is at (h, 0), as shown in *Figure 3.16*. In contrast, the graph of the function $f(x) + k = x^2 + k$ is a parabola shifted upward by k and whose vertex is at (0, k), as shown in *Figure 3.17*. Combining both horizontal and vertical shifts yields $f(x - h) + k = (x - h)^2 + k$, a parabola shifted to the right by h and upward by k and whose vertex is at (h, k), as shown in the bottom (*Figure 3.18*).

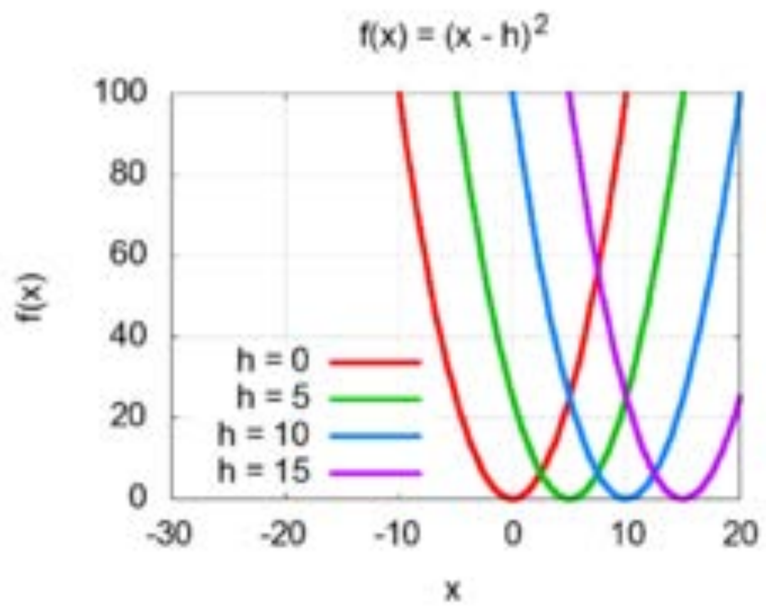


Figure 3.16 Graphs of quadratic functions shifted to the right by h = 0, 5, 10, and 15.

The graph of the function ƒ(x – h) = (x – h)2 is a parabola shifted to the right by h, whose vertex is at (h, 0).

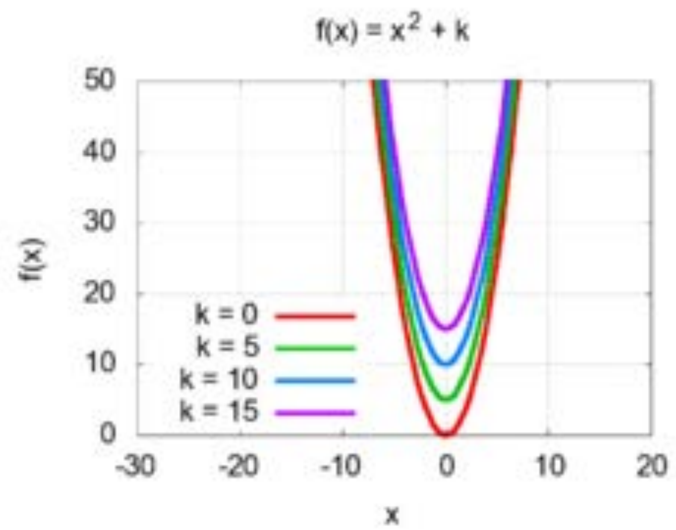


Figure 3.17 Graphs of quadratic functions shifted upward by k = 0, 5, 10, and 15.

The graph of the function ƒ(x) + k = x 2 + k is a parabola shifted upward by k, whose vertex is at (0, k).

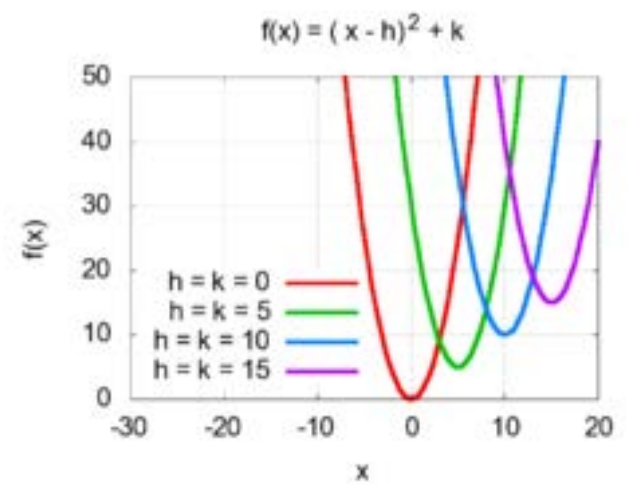


Figure 3.18 Graphs of quadratic functions shifted upward and to the right by 0, 5, 10, and 15.

The graph of the function ƒ(x – h) + k = (x – h)2 + k is a parabola shifted to the right by h and upward by k andwhose vertex is at (h, k).

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/graphs-of-quadratic-functions/quadratic-functions-of-the-form-f-x-a-x-h-2-k/

Quadratic Functions of the Type $f(x) = ax^2 + bx + c$ Where a is not Equal to 0

A quadratic function, in mathematics, is a polynomial function of the form $y = ax^2 + bx + c$.

KEY POINTS

- The graph of a quadratic function is a parabola whose axis of symmetry is parallel to the y-axis.
- If the quadratic function is set equal to zero, then the result is a quadratic equation.
- The solutions to the equation are called the roots of the equation.

The adjective quadratic comes from the Latin word quadrātum ("square"). A term like x2 is called a square in algebra because it is the area of a square with side x.

In general, a prefix quadr(i)- indicates the number 4. Examples are quadrilateral and quadrant. Quadratum is the Latin word for square because a square has four sides.

A quadratic function in the form:

$f(x) = ax^2 + bx + x$

is in standard form.

Regardless of the format, the graph of a quadratic function is a parabola (as shown in *Figure 3.19*).

- If a>0 (or is a positive number), the parabola opens upward.
- If a<0 (or is a negative number), the parabola opens downward.

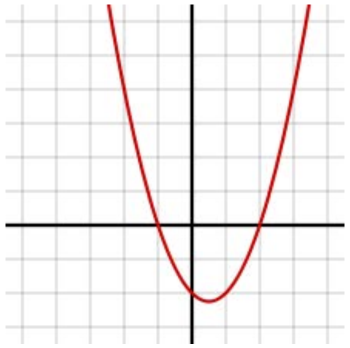


Figure 3.19 A graphed quadratic equation

where a=1, b=-1, c=-2

The coefficient a controls the speed of increase (or decrease) of the quadratic function from the vertex. A larger, positive a makes the function increase faster and the graph appear more closed.

The coefficients b and a together control the axis of symmetry of the parabola (also the x-coordinate of the vertex), which is at $x = -b/2a$.

The coefficient b alone is the declivity of the parabola as y-axis intercepts.

The coefficient c controls the height of the parabola, or more specifically, it is the point where the parabola intercepts the y-axis.

A parabola is a conic section, created from the intersection of a right circular conical surface and a plane parallel to a generating straight line of that surface. Another way to generate a parabola is to examine a point (the focus) and a line (the directrix). The locus of points in that plane that are equidistant from both the line and point is a parabola. In algebra, parabolas are frequently encountered as graphs of quadratic functions, such as

$y = x^2$

The vertex of a parabola is the place where it turns; hence, it's also called the turning point. If the quadratic function is in vertex form, the vertex is (h, k).

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/graphs-of-quadratic-functions/quadratic-functions-of-the-type-f-x-ax-2-bx-c-where-a-is-not-equal-to-0/

Applications

Graphs display complete pictures of quadratic functions and from them one can easily find critical values of the function by inspection.

KEY POINTS

- If several key points on a function are desired, it can become tedious to calculate each algebraically.
- Rather than calculating each key point of a function, one can find these values by inspection of its graph.
- Graphs of quadratic functions can be used to find key points in many different relationships, from finance to science and beyond.

Given the algebraic equation for a quadratic function, one can calculate any point on the function, including critical values like minimum/maximum and x- and y-intercepts.

These calculations can be more tedious than is necessary, however. A graph contains all the above critical points and more, and is essentially a clear and concise representation of a function. If one needs to determine several values on a quadratic function, glancing at a graph is quicker than calculating several points.

Consider the function:

$$F(x) = -\frac{x^2}{10} + 50x - 750$$

Suppose this models the profit (f(x)) in dollars that a company earns as a function of the number of products (x) of a given type that are sold, and is valid for values of x greater than or equal to 0 and less than or equal to 500.

If one wanted to find the number of sales required to break even, the maximum possible loss (and the number of sales required for this loss), and the maximum profit (and the number of sales required for this profit), one could calculate algebraically or simply reference a graph (*Figure 3.20*).

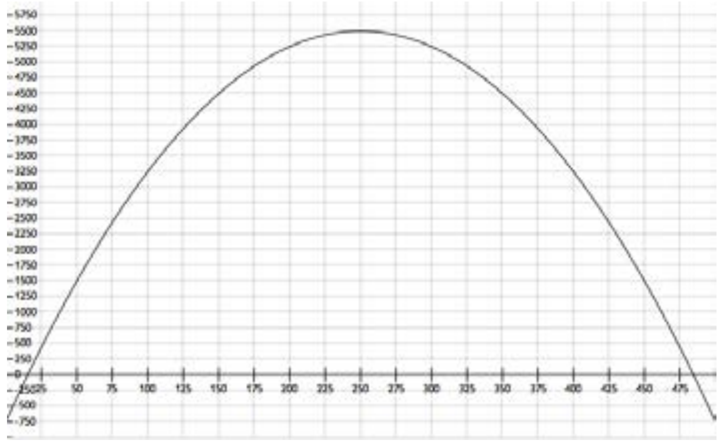


Figure 3.20 Profit versus sales

Critical points of the function can be determined by inspection. This can be less time-consuming than performing several calculations.

By inspection, we can find that the maximum loss is $750, which is lost at both 0 and 500 sales. Maximum profit is $5500, which is

achieved at 250 sales. The break-even points are between 15 and 16 sales, and between 484 and 485 sales.

The above example pertained to business sales and profits, but a similar model can be used for many other relationships in finance, science and otherwise. For example, the reproduction rate of a strand of bacteria can be modeled as a function of differing temperature or **pH** using a quadratic functionality.

Section 5

Further Equation Solving

https://www.boundless.com/algebra/functions-equations-and-inequalities/further-equation-solving/

Rational Equations

A rational equation is when two rational expressions are set equal to each other and x values are found that make the equation true.

KEY POINTS

- You can have multiple values that will satisfy the equations. It is important to check your work to make sure that all the values you find all work for the equations.
- If you have a rational equation where the denominators are the same, then the numerators must be the same.This gives us a strategy: Find a common denominator, and then set the numerators equal.
- It is important to write each equation in its simplest form. This means to factor each problem out as much as you possibly can before starting.

A rational equation means that you are setting two **rational expressions** equal to each other. The goal is to find the x value or values that make the equation true.

Suppose you are told that:

$$\frac{x}{8} = \frac{3}{8}$$

We can use simple algebra to solve this equation. There are several ways to do this. One way is to isolate the variable on one side. And because we are familiar with basic algebra, we know that what you do to one side, you must also do to the other side.

$$(\frac{8}{1}) * \frac{x}{8} = \frac{3}{8} * (\frac{8}{1})$$

The 8 on both sides will cancel out the **denominator** in the expressions and return this: $x = 3$

You can also come to this conclusion by deductive reasoning. If you think about it, the x in this equation has to be a 3. That is to say, if x=3 then this equation is true; for any other x value, this equation is false.

This leads us to a very general rule about rational equations:

If you have a rational equation where the denominators on either side of the equation are the same, then their respective **numerator**s must be the same value, even though they might be expressed differently.This suggests a strategy: Find a common denominator, and then set the numerators equal.

To start, we rewrite both fractions with their common denominator. You can find the common denominator by first factoring each expression, and then looking to see what each must be multiplied by

to be equal to one another. This is easier illustrated as an example. Lets start with this equation:

$$\frac{3}{x^2+12x+36} = \frac{4x}{x^3+4x^2-12x}$$

Now we factor the denominators:

$$\frac{3}{(x+6)^2} = \frac{4x}{x(x+6)(x-2)}$$

The next step is to make both denominators similar to each other:

$$(x)(x-2) * [\frac{3}{(x+6)^2}] = [\frac{4x}{x(x+6)(x-2)}] * (x+6)$$

$$\frac{3(x)(x-2)}{(x+6)^2(x)(x-2)} = \frac{4x(x+6)}{x(x+6)^2(x-2)}$$

Based on the rule above, since the denominators are equal, we can now assume the numerators are equal.

$$3x(x-2) = 4x(x+6)$$

We then multiply it out:

$$3x^2 - 6x = 4x^2 + 24x$$

Now, we have simplified this into a basic quadratic equation. It is always easiest to deal with equations by isolating the variables to one side:

$$0 = x^2 + 30x$$

and then factor:

$$0 = x(x+30)$$

Two solutions to the quadratic equation, x=0 and x=-30. However, in this case, x=0 is not valid, why is this? Well, lets look again at the original expression of the equation:

$$\frac{3}{x^2+12x+36} = \frac{4x}{x^3+4x^2-12x}$$

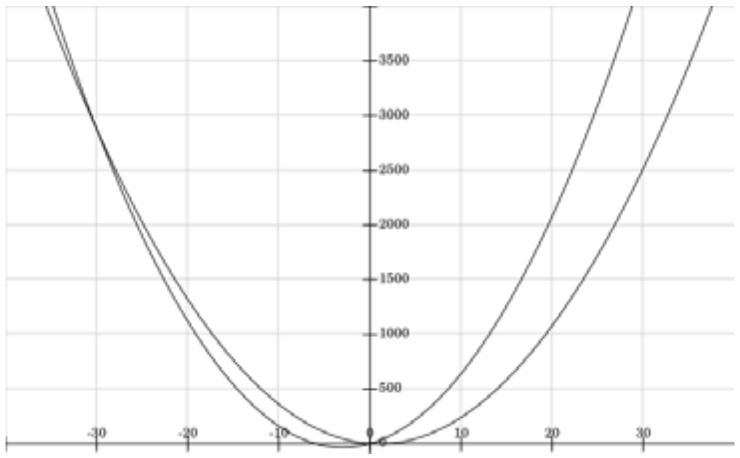
Now, lest plug in x=0:

$$\frac{3}{0^2+12*0+36} = \frac{4*0}{0^3+4*0^2-12*0}$$

$$\frac{3}{36} = \frac{0}{0}$$

We know that you can not divide a number by 0 or have any other number divide a 0, so this answer is invalid. Now, why don't we try and plug in x=-30

Figure 3.21 A graph of 3x^2-6x and 4x^2+24x



We can see that they intersect at both x=-30 and x=0. However, from our original equation, we know that x=0 is outside of the domain, and so is not a valid solution.

$$\frac{3}{(-30)^2 + 12 * (-30) + 36} = \frac{4 * (-30)}{(-30)^3 + 4 * (-30)^2 - 12 * (-30)}$$

$$\frac{3}{900 - 360 + 36} = \frac{-120}{-27000 + 3600 + 360}$$

$$\frac{3}{576} = \frac{-120}{-23040}$$

If you were to divide both the numerator and denominator by 40, you will see that the terms are equal.

So this problem actually has only one solution, x=–30.

As this example shows, it is important to always check your work.

If you graphed the two function, the two graphs would intersect at one point only: the point when x=–30.

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/further-equation-solving/rational-equations/

Radical Equations

Equations involving radicals are often solven by moving the radical to one side then squaring both sides.

KEY POINTS

- When solving equations that involve radicals, begin by asking: is there an x under the square root? The answer to this question will determine the way you approach the problem.
- If there is not an x under the square root—if only numbers are under the radicals—the problem can be solved much the same way as if it had no radicals. However, if there is an x under a square root, then move everything except the radical to one side, then square both sides.
- Squaring both sides can introduce false answers—so it is important to check the answers after solving!

When solving equations that involve **radical**s, begin by asking: is there an x under the **square root**? The answer to this question will determine the way the problem is approached.

If there is not an x under the square root—if only numbers are under the radicals—the problem can be solved much the same way as if it had no radicals.

Example 1: Radical Equation with No Variables Under Square Roots

Sample problem: no variables under radicals.

$\sqrt{2}x + 5 = 7 - \sqrt{3}x$

1. Get everything with an x on one side, everything else on the other
   $\sqrt{2}x + \sqrt{3}x = 7 - 5$
2. Factor out the x
   $x(\sqrt{2} + \sqrt{3}) = 2$
3. Then divide to solve for x
   $x = \dfrac{2}{\sqrt{2} + \sqrt{3}}$

The key thing to note about such problems is that both sides of the equation do not have to be squared. $\sqrt{2}$ may look complicated, but it is just a number—it can be found on a calculator—it functions in the equation just the way that the number 10, or $\frac{1}{3}$, or π would.

If there is an x under the square root, the problem is must be approached differently. Both sides have to be squared to get rid of

the radical. However, there are two important notes about this kind of problem:

Always get the radical alone, on one side of the equation, before squaring.

Squaring both sides can introduce false answers—so it is important to check the answers after solving!

Both of these principles are demonstrated in the following example.

EXAMPLE 2: Radical Equation with Variables under Square Roots

$\sqrt{x+2} + 3x = 5x + 1$

Here is a sample problem with variables under radicals.

$\sqrt{x+2} = 2x + 1$

1. Isolate the radical by subtracting 3x from each side. Now square both sides, giving

   $x + 2 = (2x + 1)^2$

2. Multiply out.

   $x + 2 = 4x^2 + 4x + 1$

3. Get everything on one side.

   $4x^2 + 3x - 1 = 0$

4. Factoring: the easiest way to solve quadratic equations.

   $(4x - 1)(x + 1) = 0$

   This gives two solutions, x= 1/4 and x=-1.

5. As mentioned above, check these answers to make sure they work!

   Checking x=1/4:

   $\sqrt{\frac{1}{4} + 2} + 3(\frac{1}{4}) = 5(\frac{1}{4}) + 1$

   $\sqrt{\frac{1}{4} + \frac{8}{4}} + \frac{3}{4} = \frac{5}{4} + \frac{4}{4}$

   $\sqrt{\frac{9}{4}} + \frac{3}{4} = \frac{5}{4} + \frac{4}{4}$

   $\frac{3}{2} + \frac{3}{4} = \frac{5}{4} + \frac{4}{4}$

   $\frac{9}{4} = \frac{9}{4}$

   Therefore, x=1/4 is a valid solution.

   Checking x=-1:

   $\sqrt{-1+2} + 3(-1) = 5(-1) + 1$

   $\sqrt{1} - 3 = -5 + 1$

   $1 - 3 = -5 + 1$

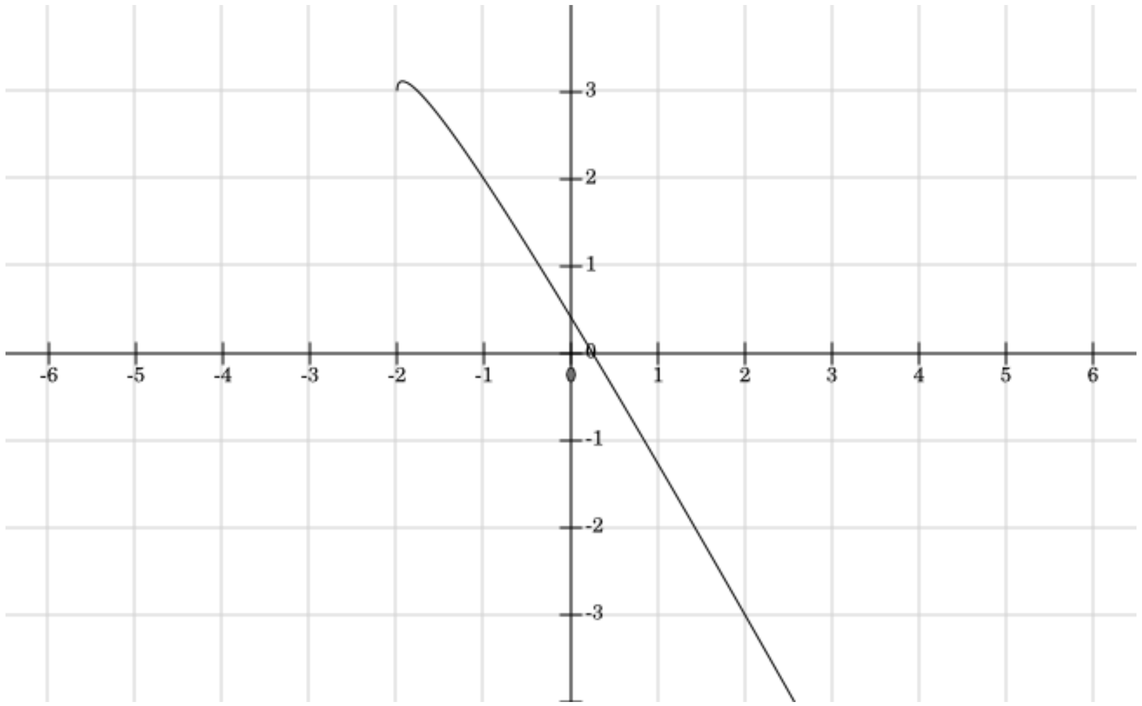
   $2 \neq -4$

   Therefore -1 is not a valid solution to this equation.

So the algebra yielded two solutions: 1/4 and –1. Checking, however, it is discovered that only the first solution is valid. This problem demonstrates how important it is to check solutions whenever squaring both sides of an equation. Looking at the equation graphically is another way to check algebraic solutions, as seen in *Figure 3.22*, it only crosses the y-axis once.

If variables under the radical occur more than once, then this procedure must be done multiple times. Each time, isolate a radical and then square both sides.

Figure 3.22 Graph of (x+2)^(1/2)-2x-1



This graph shows where the given equation is equal to zero. This is equivalent to showing where (x+2)^(1/2)+3x=5x+1.

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/further-equation-solving/radical-equations/

Equations with Absolute Value

To solve an equation with an absolute value, first move the absolute value to one side, then solve for the positive and negative cases.

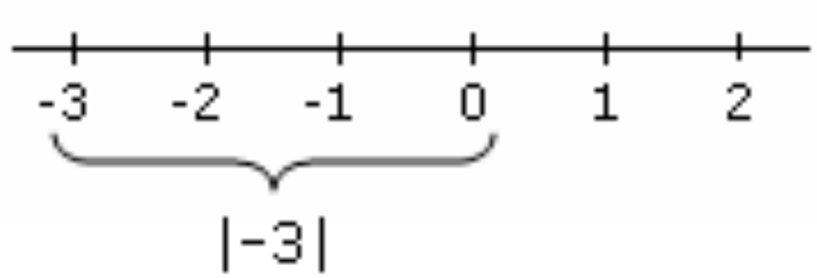
KEY POINTS

- Absolute value is one of the simplest functions—and paradoxically, one of the most problematic. On the face of it, nothing could be simpler: it just means “whatever comes in, a positive number comes out”.
- In order to solve an absolute value equation: 1) Isolate the absolute value algebraically. 2) Think through the problem. That is to say, set the absolute value term equal to the other side of the equation, then the opposite of the other side (-1 times it). 3) Finally, do more algebra to isolate x.
- Absolute values are always positive. An absolute value equation can have, at most, two solutions.

Absolute value is one of the simplest functions—and paradoxically, one of the most problematic.

At face value, nothing could be simpler: absolute value simply means “for any input, a positive value is returned.” It can also be thought of the distance a number is from 0, as shown in *Figure 3.23*.

Figure 3.23 Absolute Value



Both 3 and -3 are the same distance from 0, so the absolute value of both is 3.

$|5| = 5$ and $|-5| = 5$

But consider these three equations. They look very similar—only the number changes—but the solutions are completely different.

Equation 1: $|x| = 10$

x=10 works, as does x=-10, therefore our solution is $x = \pm 10$

Equation 2: $|x| = -10$

Now, x=10 doesn't work, and neither does x=-10. This equations has no solutions because absolute values are never negative!

Equation 3: $|x| = 0$

Now x=0 is the only solution.

We see that the first problem has two solutions, the second problem has no solutions, and the third problem has one solution. This is only one example of how absolute value equations may become confusing—and how you can solve them if you think more freely than with memorized rules.

For more complicated problems, follow a three-step approach.

1. Isolate the absolute value algebraically.
2. Think the problem through like the simpler problems above. That is to say, set the absolute value term equal to the other side of the equation, and the opposite of the other side (-1 times it)
3. Isolate the value of x.

Most problems with this type of equation do not occur in the first and third step. Also they do not occur because students try to think it through incorrectly (second step). The problems often occur because students try to take "shortcuts" to avoid the second step entirely.

Example: Solving An Absolute Value Equation (No Variable on the Other Side)

Solve the following equation for x:

$$3\left|2x+1\right|-7=5$$

1. Algebraically isolate the absolute value

   $3\left|2x+1\right|=12$

   $\left|2x+1\right|=4$

2. Think!

   For the moment, forget about the quantity 2x+1; just think of it as something. The absolute value of "something" is 4. So, in analogy to what we did before, the "something" can either be 4, or −4. So that gives us two possibilities

   $2x+1=4$

   $2x+1=-4$

3. Algebraically solve both equations for x:

   $2x=3$ or $2x=-5$

   $x=3/2$ or $x=-5/2$

   Therefore, this problem has two answers, x=3/2 and x=-5/2

Section 6

Working with Linear Inequalities

https://www.boundless.com/algebra/functions-equations-and-inequalities/working-with-linear-inequalities/

Linear Inequalities

Expressions that are designated as less than, greater than, less that or equal to or greater than or equal to, it is a linear inequality.

KEY POINTS

- When two expressions are connected by $<$, $>$, $\leq$, $\geq$ sign, we have an inequality. For inequalities that contain variable expressions, you may be asked to solve the inequality for that variable. This just means that you need to find the values of the variable that make the inequality true.
- A linear inequality is solved very similarly to how we solve equal functions. The difference the answers are:- $\geq$,contains values equal and greater than the found solution.
- Linear inequalities are commonly written out as $a_0 + a_1x_1 + a_2x_2 + \ldots + a_nx_n < 0$.

When two linear (first-order) expressions are not equal, but are designated as less than, greater than, less that or equal to or greater than or equal to, it is called a linear **inequality**. These can be denoted by, respectively: $<$, $>$, $\leq$, $\geq$. For inequalities that contain variable expressions, you may be asked to solve the inequality for that variable. This just means that you need to find the values of the variable that make the inequality true.

A linear inequality is a type of inequality which involves a linear function, or a first order function. Remember that when you solve a **linear equation**, there is usually one value that makes the equation true. But when you solve an inequality, there can be many values that make the statement true.

Look at this inequality: $x>4$.

The solution to this inequality includes every number that is greater than 4 as shown in *Figure 3.24*.

Figure 3.24 Inequality

Solutions to x>4 are graphed on the number line.

When operating in terms of **real numbers**, linear inequalities are the ones written in the forms

$$f(x) < b$$

where f(x) is a linear function in real numbers, and b is a constant real number. The above are commonly written out as

$$a_0 + a_1x_1 + a_2x_2 + \ldots + a_nx_n < 0$$

Sometimes they may be written out in the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n < b$$

Here x_1, x_2,... x_n are called the unknowns, a_0, a_1, a_2,... a_n are called the coefficients, and b is the constant term. A linear inequality looks like a linear equation, with the inequality sign replacing the equal sign.

A system of linear inequalities is a set of linear inequalities in the same variables. Not all systems of linear inequalities have solutions.

EXAMPLE

Try and solve this inequality: $7x + 3 + x \leq 1 + 4x - 10$. The first step, as in most equations, is to simplify the problem by isolating the variables to one side:

$7x + x - 4x \leq 1 - 10 - 3$
$4x \leq -12$
$x \leq -3$

This means that x is less than, or equal to -3. Meaning that: $x = -3, -4, -5, \ldots, -\infty$

Source: https://www.boundless.com/algebra/functions-equations-and-inequalities/working-with-linear-inequalities/linear-inequalities/

Compound Inequalities

Another type of inequality is the compound inequality, a compound inequality is of the form: $a < x < b$.

KEY POINTS

- Another type of inequality is the compound inequality. A compound inequality is of the form: a < x < b.
- There are two statements in a compound inequality. The first statement is a < x. The next statement is x < b. When we read this statement, we say "a is less than x," then continue saying "and x is less than b".
- An example of a compound inequality is: 4 < x < 9. In other words, x is some number strictly between 4 and 9.

Another type of **inequality** is the **compound** inequality. A compound inequality is of the form:

$$a < x < b$$

There are actually two statements here. The first statement is a < x. The next statement is x < b. When reading this statement, the forms say "a is less than x," then continue saying "and x is less than b."

Just by looking at the inequality, it can be seen that the number x is between the numbers a and b. The compound inequality a < x < b

indicates "betweenness." Without changing the meaning, the statement $a < x$ can be read as $x > a$. Thus, the form $a < x < b$ can be read as "x is greater than a and at the same time is less than b." For example:

1. $4 < x < 9$. x is some number strictly between 4 and 9. The numbers 4 and 9 are not included, so we use open circles at these points (*Figure 3.25*).
2. $2 < z < 0$. z is some number between -2 and 0.
3. $1 < x + 6 < 8$. The expression $x + 6$ represents some number strictly between 1 and 8.

Consider problem 3 above, $1 < x + 6 < 8$. The statement says that the quantity $x + 6$ is between 1 and 8. This statement will be true for only certain values of x. To solve for these values, subtract 6 from all three parts of the inequality to yield $-5 < x < 2$. Thus, if x is any number strictly between -5 and 2, the statement $1 < x + 6 < 8$ will be true.

Figure 3.25 $4 < x < 9$

Solutions to the compound inequality $4 < x < 9$ are indicated on the number line.

Inequalities with Absolute Value

Inequalities with absolute values can be solved by considering the absolute value as the distance from 0 on the number line.

KEY POINTS

- Problems involving absolute values and inequalities can be approached in at least two ways.
- Inequalities with absolute values can be solved by trial-and-error.
- Another way to solve inequalities with absolute values is to think of the absolute value as representing distance from 0 and then finding the values that satisfy that condition.

Consider the following **inequality** that includes an **absolute value**:

$|x| < 10$

Having seen that the solution to $|x|=10$ is $x=\pm10$, many students answer this question $x<\pm10$. However, this is wrong.

Here are two different, perfectly correct, ways to look at this problem.

1. What numbers work? 4 works. -4 does too. 13 doesn't work. How about -13? No: If $x=-13$, then $|x|=13$, which is not less than 10. By playing with numbers in this way, you should be able to convince yourself that the numbers that work must be somewhere between -10 and 10. This is one way to approach finding the answer.
2. The other way is to think of absolute value as representing distance from 0. $|5|$ and $|-5|$ are both 5 because both number are 5 away from 0. In this case, $|x|<10$ means "the distance between x and 0 is less than 10" - in other words, you are within 10 units of zero in either direction. Once again, we conclude that the answer must be between -10 and 10.

This answer can be visualized on the **number line** as shown in *Figure 3.26*, in which all numbers whose absolute value is less than 10 are highlighted.

Figure 3.26 Solution to $|x|<10$

All numbers whose absolute value is less than 10; –10

It is not necessary to use both of these methods; use whichever method is easier for you to understand.

More complicated absolute value problems

should be approached using the same steps as the equations discussed above: algebraically isolate the absolute value and then algebraically solve for x.

Solving Problems with Inequalities

Inequalities can be solved by similar methods to linear equations, except that multiplication by negative numbers reverses the inequality.

KEY POINTS

- Inequalities can be solved by basically the same methods as linear equations with the exception that multiplication by a negative number reverses the direction of the inequality.
- Let a, b, and c represent real numbers and assume that $a < b$. Then, if $a < b$, $a + c < b + c$ and $a - c < b - c$; if any real number is added to or subtracted from both sides of an inequality, the sense of the inequality remains unchanged.
- If c is a positive real number, then if $a < b$, $ac < bc$ and $ac < bc$. If both sides of an inequality are multiplied or divided by the same positive number, the sense of the inequality remains unchanged.
- While, if c is a negative real number, then if $a < b$, $ac > bc$ and $ac > bc$. If both sides of an inequality are multiplied or divided by the same negative number, the inequality sign must be reversed in order for the resulting inequality to be equivalent to the original inequality.

We have discovered that an equation is a mathematical way of expressing the relationship of equality between quantities. However, not all relationships need be relationships of equality. Certainly the number of human beings on earth is greater than 20. Also, the average American consumes less than 10 grams of vitamin C every day. These types of relationships are not relationships of equality but, rather, relationships of inequality.

A linear inequality is a mathematical statement that one linear expression is greater than or less than another linear expression.

The following notation is used to express relationships of inequality:

- $>$ Strictly greater than
- $<$ Strictly less than
- $\geq$ Greater than or equal to
- $\leq$ Less than or equal to

Note that the expression $x > 12$ has infinitely many solutions. Any number strictly greater than 12 will satisfy the statement. Some solutions are: 13, 15, 90, 12.1, 16.3, and 102.51.

A linear equation, we know, may have exactly one solution, infinitely many solutions, or no solution. Speculate on the number of solutions of a linear inequality. (Hint: Consider the inequalities $x < x-6$ and $x \geq 9$.)

Figure 3.27 Solving Linear Inequalities

$2x + 1 \leq 7$	Original Inequality.
$2x + 1 - 1 \leq 7 - 1$	Subtract 1 from each side to undo the addition.
$2x \leq 6$	Simplify.
$\frac{2x}{2} \leq \frac{6}{2}$	Divide each side by 2 to undo the multiplication.
$x \leq 3$	Simplify.

Miah was asked to find the values of x that make this inequality true: $2x + 1 \leq 7$.

A linear inequality may have infinitely many solutions or no solutions.

Inequalities can be solved by basically the same methods as linear equations. Let a, b, and c represent **real number**s and assume that $a < b$.

Then, if $a < b$,

- $a + c < b + c$ and $a - c < b - c$. If any real number is added to or subtracted from both sides of an inequality, the sense of the inequality remains unchanged.

- If c is a positive real number, consider then if $a < b$, $ac < bc$ and $ac < bc$. If both sides of an inequality are multiplied or divided by the same positive number, the sense of the inequality remains unchanged.
- If c is a negative real number, then if $a < b$, $ac > bc$ and $ac > bc$. If both sides of an inequality are multiplied or divided by the same negative number, the inequality sign must be reversed (change direction) in order for the resulting inequality to be equivalent to the original inequality.

Chapter 4

Polynomial and Rational Functions

https://www.boundless.com/algebra/polynomial-and-rational-functions/

Polynomial Functions and Models

https://www.boundless.com/algebra/polynomial-and-rational-functions/polynomial-functions-and-models/

The Leading-Term Test

Analysis of a polynomial reveals whether the function will increase or decrease as x approaches positive and negative infinity.

KEY POINTS

- Properties of the leading term of a polynomial reveal whether the function increases or decreases continually as x values approach positive and negative infinity.
- If n is odd and a_n is positive, the function declines to the left and inclines to the right.
- If n is odd and a_n is negative, the function inclines to the left and declines to the right.
- If n is even and a_n is positive, the function inclines both to the left and to the right.
- If n is even and a_n is negative, the function declines both to the left and to the right.

All polynomial functions of first or higher order either increase or decrease indefinitely as x values grow larger and smaller. It is possible to determine the end behavior of a polynomial function without using a graph. Consider the polynomial function:

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$$

$a_n x^n$ is called the **leading term**, while a_n is known as the **leading coefficient**. The properties of the leading term and leading coefficient indicate whether f(x) increases or decreases continually as x values approach positive and negative infinity.

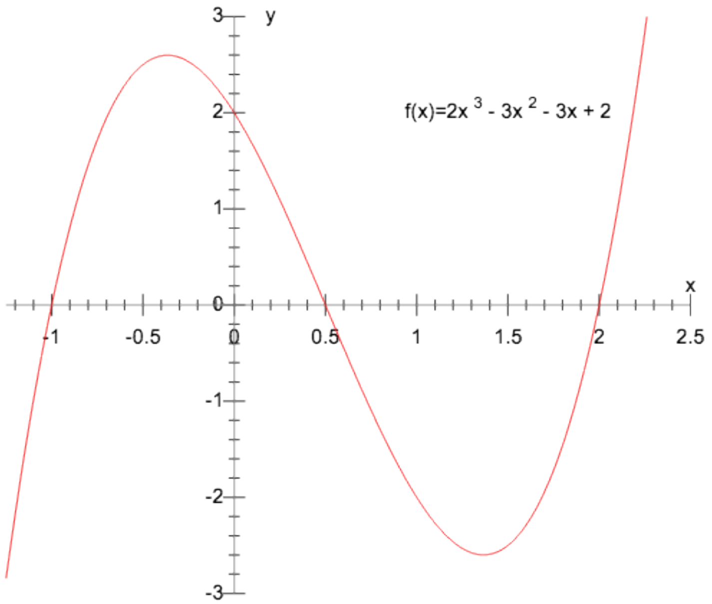
- If n is odd and a_n is positive, the function declines to the left and inclines to the right.
- If n is odd and a_n is negative, the function inclines to the left and declines to the right.
- If n is even and a_n is positive, the function inclines both to the left and to the right.
- If n is even and a_n is negative, the function declines both to the left and to the right.

Consider the polynomial:

$$f(x) = 2x^3 - 3x^2 - 3x + 2$$

In the leading term, a equals 2, and n equals 3. Because n is odd and a is positive, the graph declines to the left and inclines to the right (*Figure 4.1*).

Figure 4.1 Graph of cubic polynomial



Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/polynomial-functions-and-models/the-leading-term-test/

Finding Zeros of Factored Polynomials

The factored form of a polynomial reveals its zeros, which are defined as points where the function touches the x axis.

KEY POINTS

- A polynomial function may have zero, one, or many zeros.
- All polynomial functions of positive, odd order have at least one zero, while polynomial functions of positive, even order may not have a zero.
- Regardless of odd or even, any polynomial of positive order can have a maximum number of zeros equal to its order.

The factored form of a polynomial can reveal where the function crosses the x axis. An x value at which this occurs is called a "**zero**" or "root."

Consider the factored function:

$$f(x) = (x - a_1)(x - a_2)\dots(x - a_n)$$

Each value a_1, a_2, and so on is a zero.

A polynomial function may have zero, one, or many zeros. All polynomial functions of positive, odd order have at least one zero, while polynomial functions of positive, even order may not have a zero.

Regardless of odd or even, any polynomial of positive order can have a maximum number of zeros equal to its order. For example, a cubic function can have as many as three zeros, but no more. This is known as the fundamental theorem of algebra.

Consider the function:

$$f(x) = x^3 + 2x^2 - 5x - 6$$

This can be rewritten in the factored form:

$$f(x) = (x + 3)(x + 1)(x - 2)$$

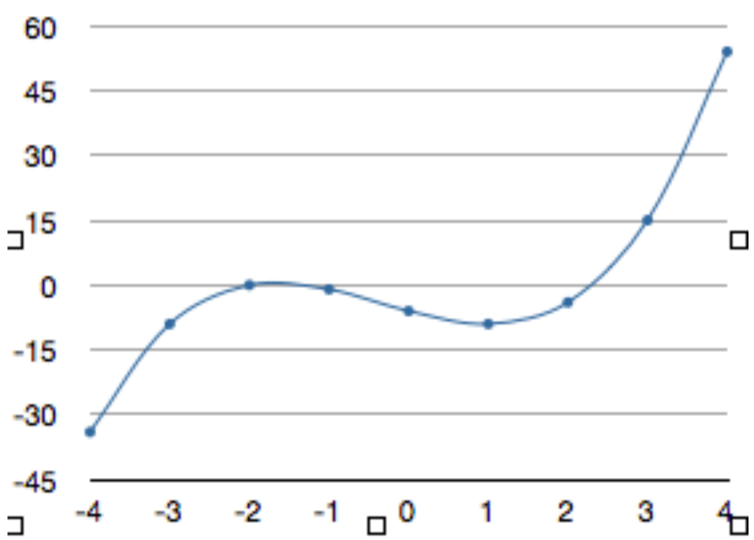


Figure 4.2 Graph of f(x)=(x+1)(x+3)(x-2)

f(x) crosses the x axis at x=-3, x=-1, and x=2.

Replacing x with a value that will make (x+3), (x+1), or (x-2) will result in f(x) being equal to zero. Thus, the zeroes for f(x) are at x=-3, x=-1, and x=2. This can also be shown graphically *Figure 4.2*.

Introduction: Polynomial and Rational Functions and Models

Functions are commonly used in fitting data to a trend line and differ in terms of accuracy and ease of use.

KEY POINTS

- Researchers will often collect many discrete samples of data, relating two or more variables, without knowing the mathematical relationship between them. Curve fitting is used to create trend lines intended to fill in the points between and beyond collected data points.
- Polynomial functions are easy to use for modeling but ill-suited to modeling asymptotes and some functional forms, and they can become very inaccurate outside the bounds of the collected data.
- Rational functions can take on a much greater range of shapes and are more accurate both inside and outside the limits of collected data than polynomial functions. However, rational functions are more difficult to use and can include undesirable asymptotes.

Polynomial and rational functions are often used in statistical modeling. For a set of data, such functions can be used to create a trend line that relates discrete findings on two or more axes.

A polynomial function has the following form:

$$y = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0$$

where n is an integer greater than or equal to 0, x and y are variables, and a_n,a_{n-1}, a_2, a_1, and a are constants.

The value of n defines the degree of the polynomial. If n=0, the function is a constant; if n=1, the function is a line; if n=2, the function is quadratic; if n=3, the function is cubic, and so on.

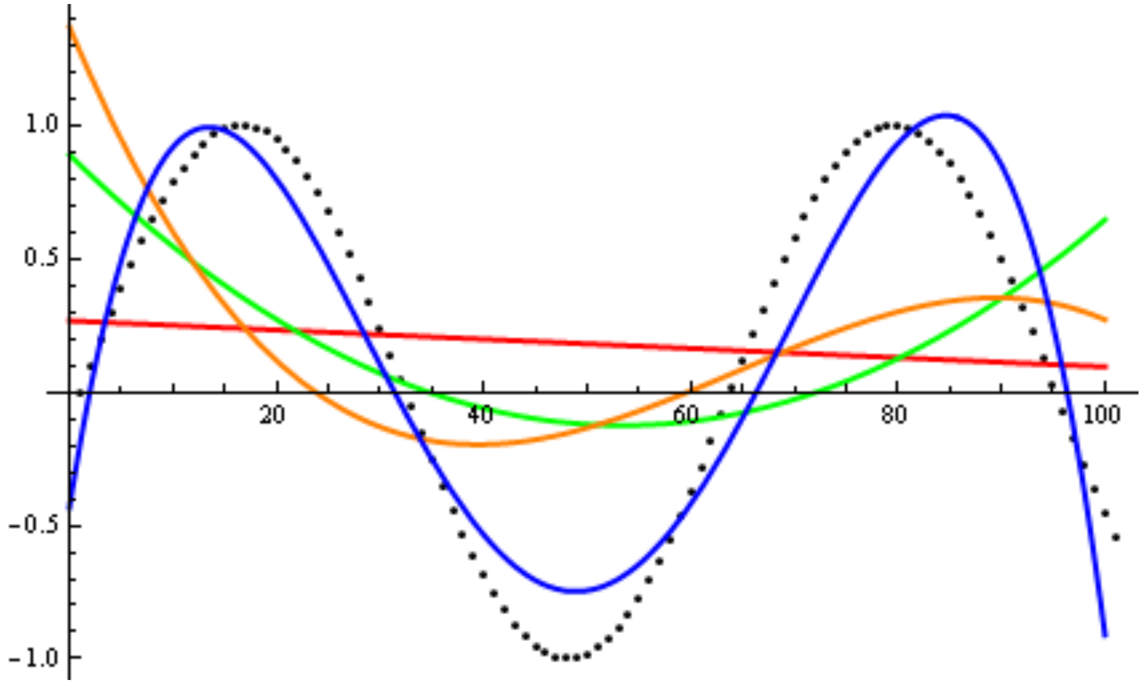
The form of the polynomial function is very versatile and can be used to represent not only complex but also simple functions. Consider the following function:

$$y = x^5 + 3x$$

in which n equals 5 and a_n equals 1. Because there are no terms for x^4, x^3, and x^2, the values of a_{n-1}, a_{n-2}, and a_2 are equal to 0. The value of a_1 is equal to 3.

Polynomial functions are very simple in form and easy to use, but they have limitations with regard to statistical modeling. They can take on only a limited number of shapes and are particularly ill-

Figure 4.3 Curve Fitting



Polynomial curves generated to fit points (black dots) of a sine function: The red line is a first degree polynomial; the green is a second degree; the orange is a third degree; and the blue is a fourth degree.

suited to modeling **asymptote**s. They also make for trend lines that become increasingly unreliable as they extend further beyond the limits of collected data.

A rational function is the ratio of two polynomial functions and has the following form:

$$y = \frac{a_n x^n + a_{n-1} + x^{n-1} + \ldots + a_2 x^2 + a_1 x + a_0}{b_m x^m + b_{m-1} + x^{m-1} + \ldots + b_2 x^2 + b_1 x + b_0}$$

Here, n and m define the degrees of the numerator and denominator, respectively, and together, they define the degree of the polynomial. For example, if n=2 and m=1, the function is described as a quadratic/linear rational function.

Rational functions are more complex in form than polynomial functions, but they have an advantage in that they can take on a much greater range of shapes and can effectively model asymptotes. They are also more accurate than polynomial functions both inside and outside the limits of collected data. However, rational functions sometimes include undesirable asymptotes that can disrupt an otherwise smooth trend line.

Section 2

Graphing Polynomial Functions

Basics of Graphing Polynomial Functions

The Intermediate Value Theorem

https://www.boundless.com/algebra/polynomial-and-rational-functions/graphing-polynomial-functions/

Basics of Graphing Polynomial Functions

A polynomial function in one real variable can be represented by a graph.

KEY POINTS

- The graph of the zero polynomial f(x) = 0 is the x-axis.
- The graph of a degree 1 polynomial (or linear function) f $(x) = a_0 + a_1 x$, where a1 ≠ 0, is an oblique line with y-intercept a0 and slope a1.
- The graph of a degree 2 polynomial $f(x) = a_0 + a_1 x + a_2 x^2$, where a2 ≠ 0 is a parabola.
- he graph of any polynomial with degree 2 or greater f $(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$, where an ≠ 0 and n ≥ 2 is a continuous non-linear curve.

In mathematics, a **polynomial** is an expression of finite length constructed from variables (also called **indeterminates**) and constants, using only the operations of addition, subtraction, multiplication, and non-negative integer exponents. However, the division by a constant is allowed, because the multiplicative inverse of a non zero constant is also a constant. For example, $x^2 - x/4 + 7$ is a polynomial, but $x^2 - 4/x + 7x^{3/2}$ is not, because its second **term** involves division by the variable x (4/x), and also because its third term contains an exponent that is not a non-negative integer (3/2). The term "polynomial" can also be used as an adjective, for quantities that can be expressed as a polynomial of some parameter, as in polynomial time, which is used in computational complexity theory.

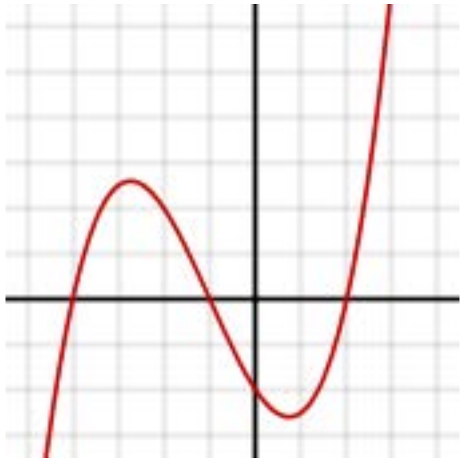


Figure 4.4 A graph of a polynomial of degree = 3

A graph of a polynomial function of degree 3

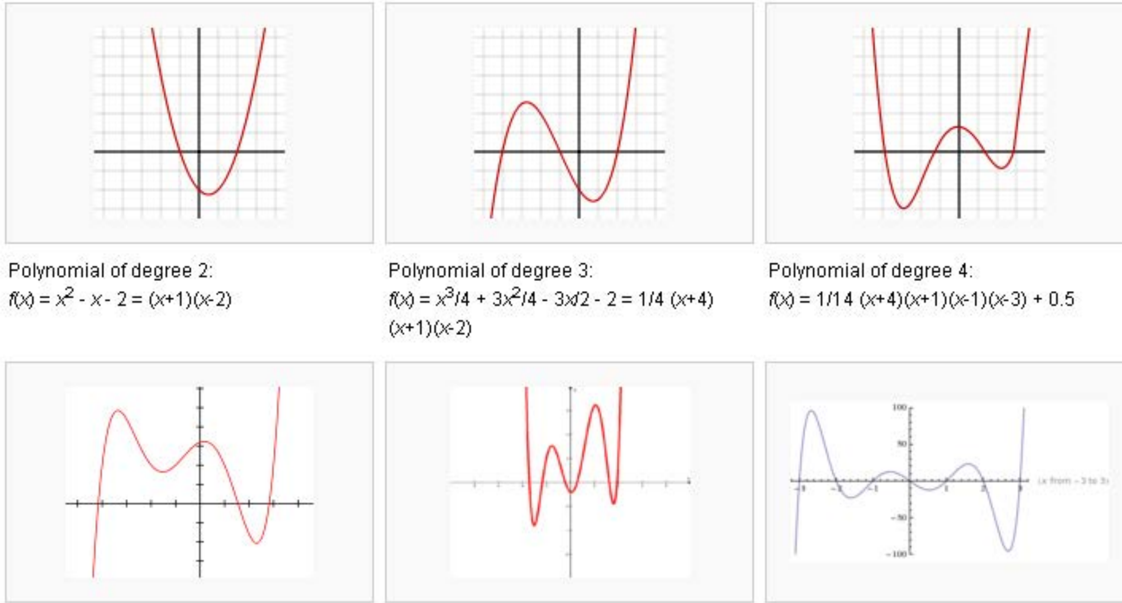
Polynomials appear in a wide variety of areas of mathematics and science. For example, they are used to form polynomial equations, which encode a wide range of problems, from elementary word problems to complicated problems in the sciences. They are used to define polynomial functions, which appear in settings ranging from basic chemistry and physics to economics and social science. They are used in calculus and numerical analysis to approximate other functions. In advanced mathematics, polynomials are used to construct polynomial rings, a central concept in abstract algebra and algebraic geometry. A typical graph of a polynomial function is shown in *Figure 4.4*.

A polynomial is either zero or can be written as the sum of a finite number of non-zero terms. Each term consists of the product of a constant (called the coefficient of the term) and a finite number of variables (usually represented by letters), also called indeterminates, raised to whole number powers. The exponent on a variable in a term is called the degree of that variable in that term; the degree of the term is the sum of the degrees of the variables in that term, and the degree of a polynomial is the largest degree of any one term. Since $x = x^1$, the degree of a variable without a written exponent is one. A term with no variables is called a constant term, or just a constant; the degree of a (non zero) constant term is 0. The coefficient of a term may be any number from a specified set. If that set is the set of real numbers, we speak of "polynomials over the reals." Other common kinds of polynomials are polynomials with integer coefficients, polynomials with complex coefficients, and polynomials with coefficients that are integers modulo of some prime number p. In most of the examples in this section, the coefficients are integers.

A polynomial function in one real variable can be represented by a graph.

- The graph of the zero polynomial $f(x) = 0$ is the x-axis.
- The graph of a degree 0 polynomial $f(x) = a_0$, where $a_0 \neq 0$, is a horizontal line with y-intercept a_0

Figure 4.5 Graphical quadratic equations



Polynomial of degree 2: $f(x) = x^2 - x - 2 = (x+1)(x-2)$

Polynomial of degree 3: $f(x) = x^3/4 + 3x^2/4 - 3x/2 - 2 = 1/4\ (x+4)(x+1)(x-2)$

Polynomial of degree 4: $f(x) = 1/14\ (x+4)(x+1)(x-1)(x-3) + 0.5$

Polynomial of degree 5: $f(x) = 1/20\ (x+4)(x+2)(x+1)(x-1)(x-3) + 2$

Polynomial of degree 6: $f(x) = 1/30\ (x+3.5)(x+2)(x+1)(x-1)(x-3)(x-4) + 2$

Polynomial of degree 7: $f(x) = (x-3)(x-2)(x-1)(x)(x+1)(x+2)(x+3)$

Quadratic equations of different orders

- The graph of a degree 1 polynomial (or linear function) $f(x) = a_0 + a_1x$, where $a_1 \neq 0$, is an oblique line with y-intercept a_0 and slope a_1.
- The graph of a degree 2 polynomial $f(x) = a_0 + a_1x + a_2x^2$, where $a_2 \neq 0$ is a parabola.
- The graph of a degree 3 polynomial $f(x) = a_0 + a_1x + a_2x^2 + a_3x^3$, where $a_3 \neq 0$, is a cubic curve.

- The graph of any polynomial with degree 2 or greater $f(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$, where $a_n \neq 0$ and $n \geq 2$ is a continuous non-linear curve.
- The graph of a non-constant (univariate) polynomial always tends to infinity when the variable increases indefinitely (in absolute value).
- Polynomial graphs are analyzed in calculus using intercepts, slopes, concavity, and end behavior.

Figure 4.5 shows a series of graphs of polynomials. Note: the degree of the polynomial in these cases can be assessed by counting the number of times the function graphically crosses the x axis.

The Intermediate Value Theorem

For each value between the bounds of a continuous function, there is at least one point where the function maps to that value.

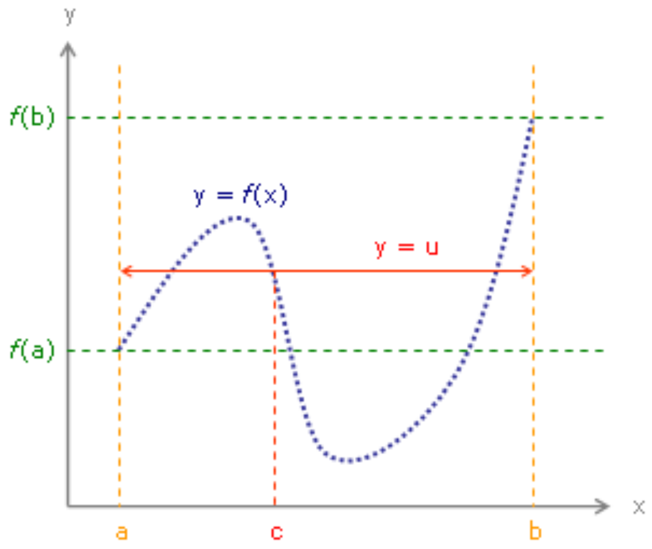
KEY POINTS

- Simply stated, the Intermediate Value Theorem points out that: if the plotted route between points A and C is smooth and continuous between point A to point C, you will have to pass through all points "B" on the journey, as long as they are on the plotted route.
- The Intermediate Value Theorem capitalizes on the completeness of functions of real numbers.
- Functions containing irrational roots do not meet the requirements of the Intermediate Value Theorem.

Simply stated, this theorem points out that, if the plotted route between points A and C is smooth and **continuous** between point A to point C, you will have to pass through all points "B" on the journey, as long as they are on the plotted route. In a clearer example, when your car accelerates from 0 to 100 mph, at some point, even for a microsecond, your car is traveling at 21.5 mph.

When you bike between points X and Z, and your path follows a semicircular route, you will bike through any point on a semicircle connecting points X and Z.

Figure 4.6 The Intermediate Value Theorem



In plotting a continuous and smooth function between two points, all points on the function between the extremes are described and predicted by the Intermediate Value Theorem.

Stated in the language of algebra supported by *Figure 4.6*: If f is a real-valued continuous function on the **interval** [a, b], and u is a number between f(a) and f(b), then there is a $c \in [a, b]$ such that f(c) = u. This can be clearly seen by following the curve from point b.

It is frequently stated in the following equivalent form: Suppose that $f : [a, b] \rightarrow R$ is continuous and that u is a real number satisfying $f(a) < u < f(b)$ or $f(a) > u > f(b)$. Then for some $c \in [a, b]$, f(c) = u. In other words, regardless of whether f(a)>f(b) or f(b)>f(a), f(c) lies between them.

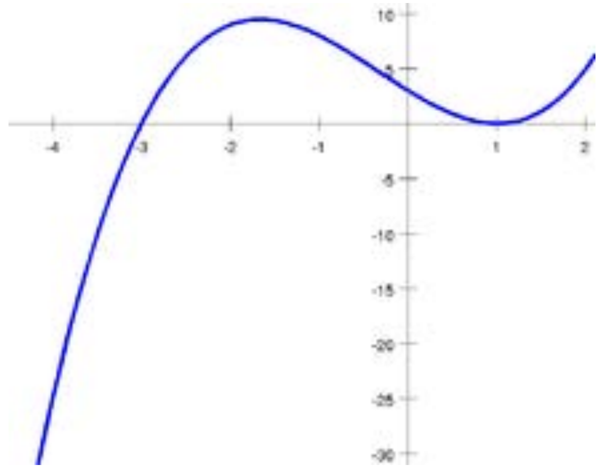


Figure 4.7 A Continuous Function

A graphed third-order equation meeting the requirements of the Intermediate Value Theorem

The theorem depends on (and is actually equivalent to) the completeness of the real numbers. It is false for the rational numbers Q. For example, the function $f(x) = x^2 - 2$ for $x \in Q$ satisfies $f(0) = -2$ and $f(2) = 2$. However there is no rational number x such that f(x) = 0, because $\sqrt{2}$ is irrational.

All polynomial functions, such as *Figure 4.7*

$$f(x) = x^3 + x^2 - 5x + 11$$

are continuous - there are no singularities or discontinuities. However, in equations of the form

$$f(x) = \frac{(2x - 1)}{(x + 2)}$$

a problem occurs when x=-2, as seen in *Figure 4.8*. The function is defined for all real numbers $x \neq -2$ and is continuous at every such point. The question of continuity at $x = -2$ does not arise, since x =

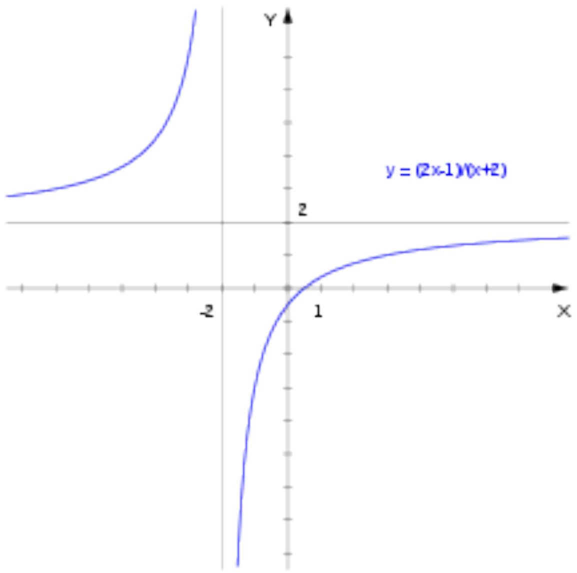


Figure 4.8
Graphing a rational function

A discontinuity occurs when x = -2: the function is not defined at x=-2.

-2 is not in the domain of f. There is no continuous function $F: R \to R$ that agrees with $f(x)$ for all $x \neq -2$.

If f is continuous on [a, b] and f(a) and f(b) differ in sign, then, at some point c in [a, b], f(c) must equal zero.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/graphing-polynomial-functions/the-intermediate-value-theorem/

Polynomial Division; The Remainder and Factor Theorems

https://www.boundless.com/algebra/polynomial-and-rational-functions/polynomial-division-the-remainder-and-factor-theorems/

Division and Factors

Polynomial long division functions similarly to long division, and if the division leaves no remainder, then the divisor is called a factor.

KEY POINTS

- Dividing one polynomial by another can be achieved by using long division. The rules for polynomial long division are the same as the rules learned for long division of integers.
- The four steps of long division are divide, multiply, subtract, and bring down.
- After completing polynomial long division, it is good to check the answers, either by plugging in a number or by multiplying the quotient times the divisor to get the dividend back.

Simplifying, multiplying, dividing, adding, and subtracting rational expressions are all based on the basic skills of working with fractions. Dividing polynomials is based on an even earlier skill, one that pretty much everyone remembers with horror: long division.

To refresh one's memory, try dividing $\frac{745}{3}$ by hand. The answer should end up as something that looks something like *Figure 4.9*.

Therefore, it can be concluded that 745/3 is 248 with a remainder of 1, or, to put it another way, 745/3 =248 + 1/3.

Long division is a skill that many may have decided they could forget, since calculators perform this task much faster. However, long division comes roaring back, because here is a problem that a calculator cannot solve: $\frac{6x^3-8x^2+4x-2}{2x-4}$. This problem is solved in a very similar way as the previous problem.

Start by rewriting the problem in standard long division form (A). Follow along the text with the graphic *Figure 4.10*.

Then divide the first time to get $3x^2$ (B). Why $3x^2$? This comes from the question: "How many times does $2x$ go into $6x^3$" or, to put the same question another way: "What would I multiply $2x$ by, in order to get $6x^3$?" This is comparable to the first step in the long division problem: "What do I multiply 3 by, to get 7?"

Now (C), multiply the $3x^2$ times the $(2x-4)$ and get $6x^3-12x^2$. Then subtract this from the line above it. The $6x^3$ terms cancel—that shows the right term was picked above. Note that careful work must

Figure 4.9 745 divided by 3

```
   248 r1
3 | 745
    6
    14
    12
     25
     24
      1
```

The long division is shown here explicitly to serve as a refresher for more complicated long division of polynomials.

be done with the signs, $-8x^2-(-12x^2)$ gives positive $4x^2$.

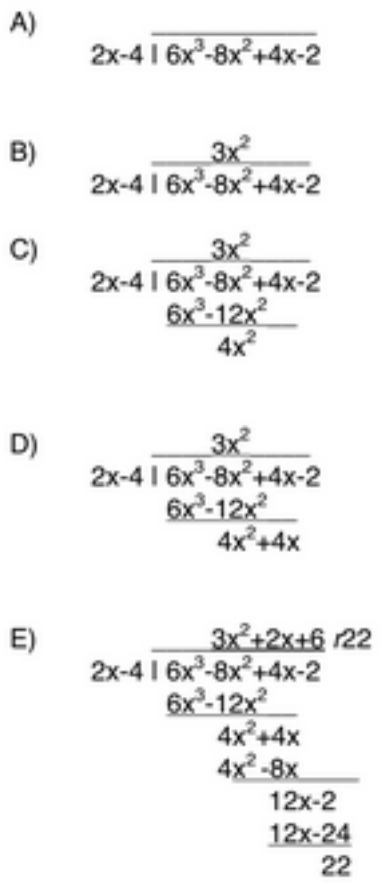
Next, bring down the $4x$, as shown in (D). All four steps of long division are now complete—divide, multiply, subtract, and bring down. At this point, the process begins again, with the question "How many times does $2x$ go into $4x^2$?"

(E) is not the next step. This is merely what the process looks like after all the steps have been finished. It is a good idea to go through the problem once more to check the work.

Therefore, it can be concluded that $\dfrac{6x^3-8x^2+4x-2}{2x-4}$ is $3x^2+2x+6$ with a remainder of 22, or, to put it another way, $3x2+2x+6+\dfrac{22}{2x-4}$.

If a polynomial can be divided by another equation and have no remainder, then the equation that was divided by is called a factor. In this case, a remainder will not be written, as the **divisor** divided evenly into the **dividend**.

Figure 4.10 Polynomial long divion



For explanations of each step, see the text.

Be sure to check the answers after doing these types of problems, either by plugging in numbers, or by multiplying the divisor by the **quotient** to see if the dividend can be gotten back!

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/polynomial-division-the-remainder-and-factor-theorems/division-and-factors/

The Remainder Theorem and Synthetic Division

Synthetic division is a technique for dividing a polynomial and finding the quotient and remainder.

KEY POINTS

- Synthetic division is most commonly applied when dividing by a monomial such as x-a.
- The most useful aspects of synthetic division are that it allows one to calculate without writing variables and uses fewer calculations.
- In algebra, synthetic division is a method of performing polynomial long division, with less writing and fewer calculations.

In algebra, the **polynomial remainder** theorem or little Bézout's theorem is an application of polynomial long division. It states that the remainder of a polynomial $f(x)$ divided by a linear divisor $(x - a)$ is equal to $f(a)$.

For example, take $f(x) = x^3 - 12x^2 - 42$ and divide it by $x - 3$. This gives the quotient $x^2 - 9x - 27$ and the remainder -123. Therefore, $f(3) = -123$. We can check this by plugging 3 into the equation above, which gives $3^3 - 12(3)^2 - 42$:

$27 - 12(9) - 42$

$27 - 108 - 42$

-123

In order to find the remainder so as to use the remainder theorem, one must first perform division. One way of doing polynomial division is to use synthetic division. In algebra, synthetic division is a method of performing polynomial long division, with less writing and fewer calculations. It is mostly taught for division by binomials of the form $x - a$, but the method generalizes to division by any monic polynomial. The most useful aspects of synthetic division are that it allows one to calculate without writing variables and uses fewer calculations. As well, it takes significantly less space than long division. Most importantly, the subtractions in long division are converted to additions by switching the signs at the very beginning, preventing sign errors. Synthetic division for linear denominators is also called division through Ruffini's rule.

The first example is synthetic division with only a monic linear denominator $x - a$:

$$\frac{x^3 - 12x^2 - 42}{x - 3}$$

The steps of synthetic division are outlined in *Figure 4.11*. The text for each step follows.

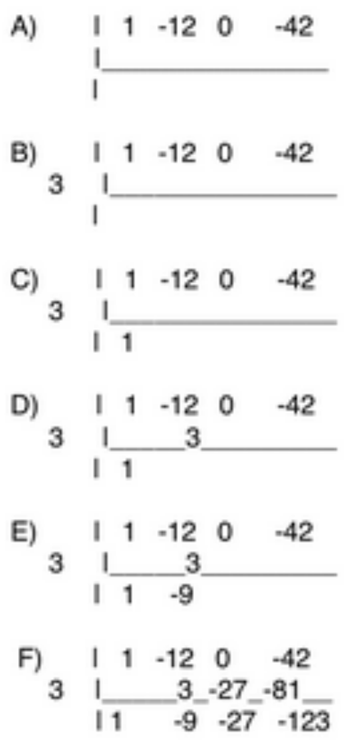


Figure 4.11 The steps of synthetic division

Follow along each step in the text.

A. Write the coefficients of the polynomial to be divided at the top (the zero is for the unseen 0x). Next negate the coefficients of the divisor: $-1x+3$.

B. Write in every coefficient of the divisor but the first one on the left.

C. Note the change of sign from –3 to 3. "Drop" the first coefficient after the bar to the last row.

D. Multiply the dropped number by the number before the bar, and place it in the next column.

E. Perform an addition in the next column.

F. Repeat the previous two steps to obtain the values shown.

Count the terms to the left of the bar. Since there is only one, the remainder has degree zero. Mark the separation with a vertical bar, as shown:

1 -9 -27 | -123

The terms are written with increasing degree from right to left beginning with degree zero for both the remainder and the result.

$$\frac{x^3-12x^2-42}{x-3}=x^2-9x-27-\frac{123}{x-3}$$

Finding Factors of Polynomials

Finding factors of polynomials is important, since it is always best to work with the simplest version of a polynomial.

KEY POINTS

- Factoring is a critical skill in simplifying functions and solving equations.
- There are four types of factoring shown which are "pulling out" common factors, factoring perfect squares, the difference between two squares, and then how to factor when the other three techniques are not applicable.
- The first step should always be "pulling out" common factors. Even if this does not factor out the polynomial completely, this will make the rest of the process much easier.

When using multiplication things are put together. When **factoring**, things are pulled apart. Factoring is a critical skill in simplifying functions and solving equations.

There are four basic types of factoring. In each case, it is beneficial to start by showing a multiplication problem—then show how to use factoring to reverse the results of that multiplication.

"Pulling Out" Common Factors

This type of factoring is based on the distributive property, which states: $2x(4x^2 - 7x + 3) = 8x^3 - 14x^2 + 6x$

When factoring, this property is done in reverse. Therefore, starting with an expression such as the one above it can be noted that every one of those terms is divisible by 2. Also, every one of those terms is divisible by x. Hence, one can "factor out," or "pull out," 2x.

$8x^3 - 14x^2 + 6x = 2x(? - ? + ?)$

For each term, it can be shown what happens when that term is divided by 2x. For instance, if 8x3 is divided by 2x, then the answer is 4x2. Doing this process for each term, the result is:

$8x^3 - 14x^2 + 6x = 2x(4x^2 - 7x + 3)$

For many types of problems, it is easier to work with this factored form.

As another example, consider 6x+3. The **common factor** is 3. When factoring 3 from 6x, 2x is left. When factoring 3 out of the 3, only 1 remains.

$6x + 3 = 3(2x + 1)$

There are two key points to understand about this kind of factoring:

1. This is the simplest kind of factoring. Whenever trying to factor a complicated expression, always begin by looking for common factors that can be pulled out.
2. The factor must be common to all the terms. For instance, $8x^3-14x^2+6x+7$ has no common factor, since the last term, 7, is not divisible by 2 or x.

Factoring Perfect Squares

The second type of factoring is based on the "squaring" formulae:

$$(x+a)^2 = x^2 + 2ax + a^2$$

$$(x-a)^2 = x^2 - 2ax + a^2$$

For instance, if the problem is x^2+6x+9, then one may recognize the signature of the first formula: the middle term is three doubled, and the last term is three squared. Thus, this simplifies to $(x+3)^2$. Attempt to notice patterns, and the problems will become easier.

$$x^2 + 10x + 25 = (x+5)^2$$

$$x^2 + 2x + 1 = (x+1)^2$$

If the middle term is negative, then the second formula is:

$$x^2 - 8x + 16 = (x-4)^2$$

$$x^2 - 14x = 49 = (x-7)^2$$

This type of factoring only works in this specific case: the middle number is something doubled, and the last number is that same value squared. Furthermore, although the middle term can be either positive or negative, the last term cannot be negative. This is because if a negative is squared, the answer is positive.

The Difference Between Two Squares

The third type of factoring is based on the third of the basic formulae:

$$(x+a)(x-a) = x^2 - a^2$$

This formula can be run in reverse whenever subtracting two perfect squares. For instance, if there is x^2-25, it can be seen that both x^2 and 25 are perfect squares. Therefore it can factor as $(x+5)(x-5)$. Other examples include:

Example 1

$$x^2 - 64 = (x+8)(x-8)$$

$$16y^2 - 49 = (4y+7)(4y-7)$$

$$2x^2 - 18 = 2(x^2-9) = 2(x+3)(x-3)$$

Note that, in the last example, the first step is done by pulling out a 2, and there are two perfect squares left. This follows the rule: always begin by pulling out common factors before trying anything else.

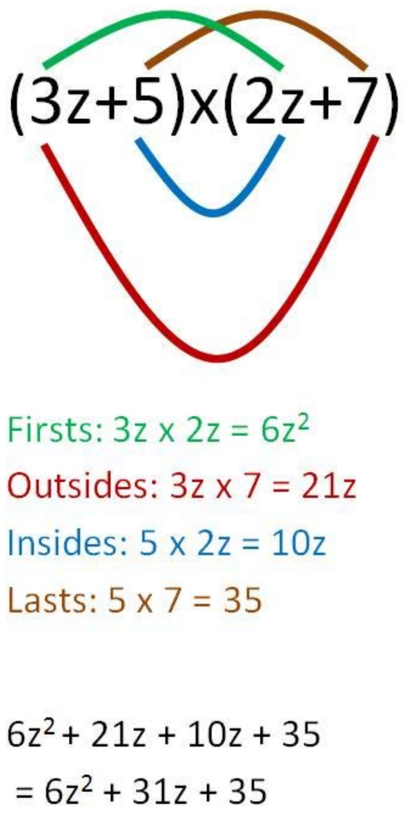


Figure 4.12 FOIL Method Diagram

start by multiplying the First terms, then the Outside terms, then the Inside terms, and finally the Last terms. Often, the outside and inside terms can eventually be added together. It is important to understand this method, in order to be able to perform it in reverse.

It is also important to note that the sum of two squares cannot be factored. x^2+4 is a perfectly good function, but it cannot be factored.

Brute Force, Old-Fashioned, Bare-Knuckle, No-Holds-Barred Factoring

In this case, the multiplication that is being reversed is just FOIL, a mnemonic for multiplying two binomials, reminding: First, Outside, Inside, Last, as shown here: *Figure 4.12*. For instance, consider:

$$(x+3)(x+7) = x^2 + 3x + 7x + 21$$

The 3 and 7 added to yield the middle term,10, and multiplied to yield the final term, 21. This can be generalized as:

Therefore, if you are given a problem such as the one above to factor, look for two numbers that add up to 10, and multiply to 21.There are a lot of pairs of numbers that add up to 10, but relatively few that multiply to 21, so start by looking for factors of 21.

The example sections provides more information.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/polynomial-division-the-remainder-and-factor-theorems/finding-factors-of-polynomials/

Section 4

Zeroes of Polynomial Functions and Their Theorems

https://www.boundless.com/algebra/polynomial-and-rational-functions/zeroes-of-polynomial-functions-and-their-theorems/

The Fundamental Theorem of Algebra

The fundamental theorem states that every non-constant single-variable polynomial with complex coefficients has at least one complex root.

KEY POINTS

- The fundamental theorem of algebra states that every non-constant single-variable polynomial with complex coefficients has at least one complex root. This includes polynomials with real coefficients, since every real number is a complex number with zero imaginary part.
- Equivalently (by definition), the fundamental theorem states that the field of complex numbers is algebraically closed.
- The fundamental theorem is also stated as follows: every non-zero, single-variable, degree n polynomial with complex coefficients has, counted with multiplicity, exactly n roots. The equivalence of the two statements can be proven through the use of successive polynomial division.

The fundamental theorem of algebra states that every non-constant single-variable polynomial with complex coefficients has at least one complex root. This includes polynomials with real coefficients, since every real number is a complex number with zero imaginary part. Equivalently (by definition), the theorem states that the field of complex numbers is algebraically closed. The theorem is also stated as follows: every non-zero, single-variable, degree n polynomial with complex coefficients has, counted with **multiplicity**, exactly n roots. The equivalence of the two statements can be proven through the use of successive polynomial division.

In spite of its name, there is no purely algebraic proof of the theorem, since any proof must use the completeness of the reals (or some other equivalent formulation of completeness), which is not an algebraic concept. Additionally, it is not fundamental for modern algebra; its name was given at a time when the study of algebra was mainly concerned with the solutions of polynomial equations with real or complex coefficients.

Complex-Analytic Proof

Find a closed disk D of radius r centered at the origin such that $|p(z)| > |p(0)|$ whenever $|z| \geq r$. The minimum of $|p(z)|$ on D, which must exist since D is compact, is therefore achieved at some point z_0 in the interior of D, but not at any point of its boundary. The Maximum modulus principle (applied to $1/p(z)$) implies then that $p(z_0) = 0$. In other words, z_0 is a zero of $p(z)$.

Figure 4.13 The Maximum Modulus Principle

A plot of the modulus of cos(z) (in red) for z in the unit disk centered at the origin (shown in blue). As predicted by the fundamental theorem, the maximum of the modulus cannot be inside of the disk (so the highest value on the red surface is somewhere along its edge).

In mathematics, the maximum modulus principle *Figure 4.13* in complex analysis states that if f is a holomorphic function, then the modulus cannot exhibit a true local maximum that is properly within the domain of f. In other words, either f is a constant function, or, for any point z_0 inside the domain of f there exist other points arbitrarily close to z_0 at which $|f|$ takes larger values.

Finding Polynomials with Given Zeroes

To construct a polynomial from given zeroes, set x equal to each zero, move everything to one side, then multiply each resulting equation.

KEY POINTS

- A polynomial constructed from n roots will have degree n or less. That is to say, if given three roots, then the highest exponential term needed will be x^3.
- Each zero given will end up being one term of the factored polynomial. After finding all the factored terms, simply multiply them together to obtain the whole polynomial.
- Because a polynomial and a polynomial multiplied by a constant have the came roots, every a polynomial is constructed from given zeroes the general solution includes a constant, shown here as c.

One type of problem is to generate a **polynomial** from given **zeroes**. Several provisions are worth noting during the excursion into this territory. Remember that the degree of a polynomial, the highest exponent, dictates the maximum number of roots it can have. Therefore, if given two zeroes then a polynomial of second

degree only needs constructed, that is, one with highest exponential term x^2.

One thing to remember, when starting with the zeroes and creating a polynomial from them, an unknown constant must always be added. For example, if given x=a and x=b as constants, then the resulting initial terms would be a constant c times the two equation that give zeroes at the appropriate place:

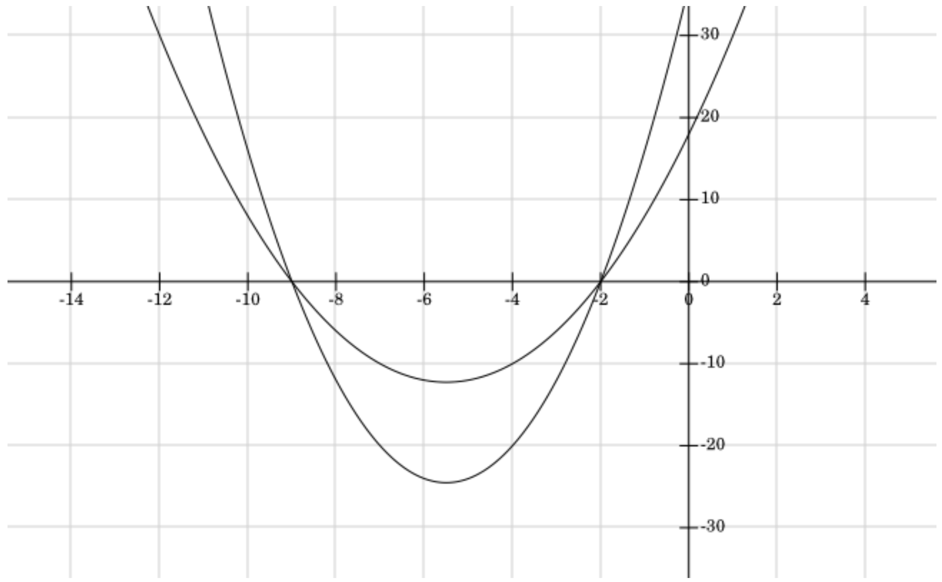
$c(x - a)(x - b)$

Multiplied out, this gives:

$cx^2 - c(a + b)x + abc.$

As an example, find a polynomial with zeroes x=-2 and x=-9. Simply manipulating these equations so that all terms are on one side of the equation, the results are x+2=0 and x+9=0. These two terms make up the factored polynomial, giving c(x+2)(x+9)=0. Notice that a c term as a constant is included, as any constant multiplied by this equation will still give zero at the appropriate places. Multiplying out, the result is $c(x^2 + 11x + 18)$. To see the effect of this constant more clearly, look at the graph of $(x^2 + 11x + 18)$ and $2(x^2 + 11x + 18)$ (*Figure 4.14*). Notice how the intercepts do not change, even though it has been multiplied by a different constant. This is because any number times 0 still equals 0.

Figure 4.14 A graph of (x2+11x+18) and 2(x2+11x+18).



Notice how the intercepts do not change, even when we multiply the function by a constant.

As another example, find a polynomial with given zeroes x=1 and x=12. Again, manipulate each of these equations until zero is on one side, giving x-1=0 and x-12=0. Entering these two conditions, the result is c(x-1)(x-12) as the factored polynomial. Multiplying out, the answer is $c(x^2 - 13x + 12)$.

Zeroes of Polynomial Functions with Real Coefficients

A root, or zero, of a polynomial function is a value that can be plugged into the function and yield zero.

KEY POINTS

- Real numbers include all the rational and irrational numbers. For example: $-5, 4/3$, $\sqrt{2}$ are all real numbers.
- If given the function, $f(x) = 0$ is a root of the function. For this reason, roots are often referred to as a zero of the function.
- There are many ways to find the roots of a polynomial. If one is confident factoring out polynomials into their simplest forms, its roots can usually be found by inspection. However, if one is not confident, or it is a tricky polynomial, the quadratic equation can be used.

The zero of a function, $f(x)$, refers to the value or values of x that will result in the function equaling zero, $f(x) = 0$. These are often called the **roots** of the function. There are many methods to find the roots of a function. One can simply factor out many functions and determine by inspecting their roots. For example, if given the function: $f(x) = x^2 - x - 2$, one can factor that out: $f(x) = (x + 1)(x - 2)$. By using basic algebraic knowledge, it is known that if one uses $f(-1)$ or $f(2)$ the result will be 0. Therefore both -1 and 2 are roots of the function.

This section specifically deals with polynomials that have real coefficients. A **real number** is any rational or irrational number, such as -5, 4/3, or even √2. An example of a non-real number would be √-1.

Even though all polynomials have roots, not all roots are real numbers. Some roots can be complex, but no matter how many of the roots are real or complex, there are always as many roots as there are powers in the function.

This is demonstrated with a quadratic equation in *Figure 4.15*.

While there are many types of polynomials and many ways to find their roots, one type of method is explored below. As a reminder, a quadratic function has the following form:

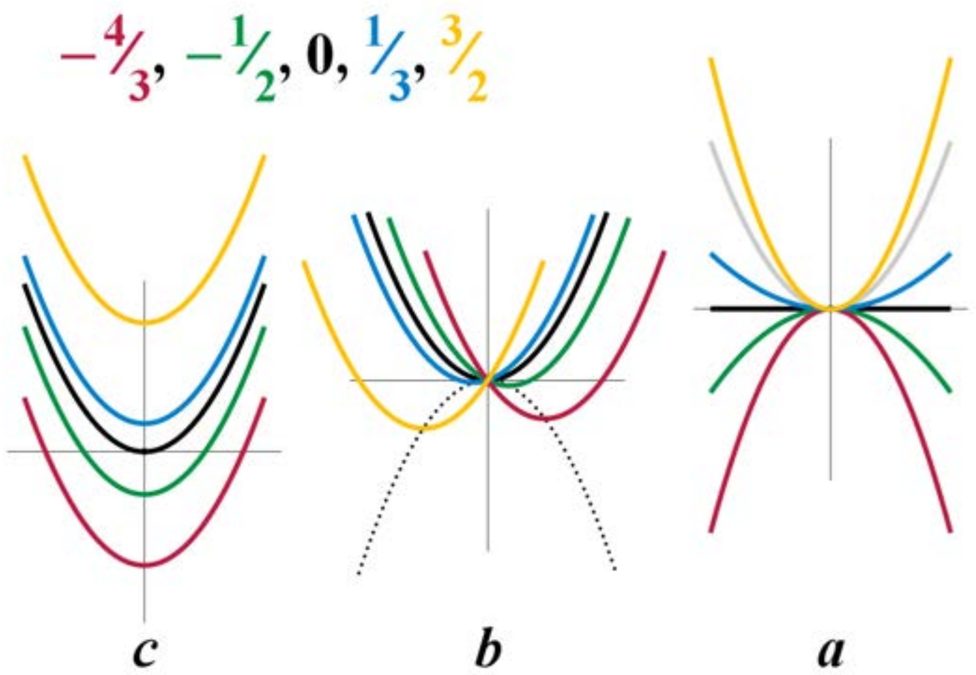
$$ax^2 + bx + c$$

This type of polynomial can be solved, meaning its roots can be found, using this equation:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

More complicated equations also exist for the higher functions, such as cubic and quartic functions, though their expressions are beyond the scope of this atom.

Figure 4.15 Plots of Quadratic Equations



Plots of real-valued quadratic function ax2 + bx + c, varying each coefficient separately.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/zeroes-of-polynomial-functions-and-their-theorems/zeroes-of-polynomial-functions-with-real-coefficients/

Rational Coefficients

Polynomials with rational coefficients should be treated and worked the same as other polynomials.

KEY POINTS

- In mathematics, a rational number is any number that can be expressed as the quotient or fraction p/q of two integers, with the denominator q not equal to zero.
- A real number that is not rational is called irrational. Irrational numbers include √2, π, and e.
- Polynomials with rational coefficients can be treated just like any other polynomial, just remember to utilize all the properties of fractions necessary during your operations.

In mathematics, a rational number is any number that can be expressed as the **quotient** or fraction p/q of two integers, with the denominator q not equal to zero. Since q may be equal to 1, every integer is a rational number. The set of all rational numbers is usually denoted by a boldface Q (or Unicode ℚ). It was thus named in 1895 by Peano after quoziente, Italian for "quotient".

The decimal expansion of a rational number always either terminates after a finite number of digits or begins to repeat the same finite sequence of digits over and over. Moreover, any

repeating or terminating decimal represents a rational number. These statements hold true not just for base 10, but also for binary, hexadecimal, or any other integer base.

A real number that is not rational is called irrational. **Irrational numbers** include √2, π, and e. The decimal expansion of an irrational number continues forever without repeating. Since the set of rational numbers is countable, and the set of real numbers is uncountable, almost all real numbers are irrational.

Zero divided by any other integer equals zero. Therefore zero is a rational number, but division by zero is undefined.

The term rational in reference to the set Q refers to the fact that a rational number represents a ratio of two integers. In mathematics, the adjective rational often means that the underlying field considered is the field Q of rational numbers. Rational polynomial usually, and most correctly, means a polynomial with rational coefficients, also called a "polynomial over the rationals". However, rational function does not mean the underlying field is the rational numbers, and a rational algebraic curve is not an algebraic curve with rational coefficients.

Finding Zeroes of a Polynomial with Rational Coefficients

Polynomials with rational coefficients can be treated just like any other polynomial, just remember to utilize all the properties of

Figure 4.16 Graph of (2x^2)/9+7x/3+6

We can graph this equation, and in doing so see where it intercepts the y axis, as a means of checking our solutions to this problem.

fractions necessary during your operations. Multiplying fractions a/b times c/d gives (ac)/(bd), whereas if one wanted to add a/b plus c/d, first convert them into ad/bd and cb/db, giving (ad+cb)/(db).

For example, the polynomial $\frac{2x^2}{9}+\frac{7x}{3}+6$ can be factored to give $(\frac{x}{3}+2)(\frac{2x}{3}+3)$. By setting each term to zero, it can be found that the zeros for this equation are x=-6 and x=-9/2. This matches what is observed graphically, as shown in *Figure 4.16*.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/zeroes-of-polynomial-functions-and-their-theorems/rational-coefficients/

Integer Coefficients and the Rational Zeroes Theorem

Each solution to a polynomial, expressed as x=p/q, which satisfies p and q are integer factors of a_0 and a_n, respectively.

KEY POINTS

- In algebra, the Rational Zeros Theorem (also known as Rational Root Theorem, or Rational Root Test) states a constraint on rational solutions (or roots) of the polynomial equation $a_nx^n + a_{n-1}x^{n-1} + \ldots + a_0 = 0$ with integer coefficients.
- If a_0 and a_n are non-zero, then each rational solution x, when written as a fraction x = p/q in lowest terms (i.e., the greatest common divisor of p and q is 1), satisfies 1) p is an integer factor of the constant term a_0, and 2) q is an integer factor of the leading coefficient a_n.
- A proof can be derived by first moving the constants to one side, factoring and multiplying by q^n. Then a generalized form of Euclid's lemma states that p divides a_0. The proof for q is similar.

In algebra, the Rational Zeros Theorem (or Rational Root Theorem, or Rational Root Test) states a constraint on rational solutions (also known as zeros, or roots) of the polynomial equation:

$$a_nx^n + a_{n-1}x^{n-1} + \ldots + a_0 = 0$$

With integer coefficients $a_n, a_{n-1}, \ldots a_0$.

If a_0 and a_n are nonzero, then each rational solution x, when written as a fraction x = p/q in lowest terms, satisfies:

- p is an integer factor of the constant term a_0.
- q is an integer factor of the leading coefficient a_n.

An Elementary Proof

Let $P(x) = a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0$ for some $a_0, \ldots, a_n \in Z$, and suppose $P(p/q) = 0$ for some **coprime** $p, q \in Z$:

$$P(\frac{p}{q}) = a_n(\frac{p}{q})^n + a_{n-1}(\frac{p}{q})^{n-1} + \ldots + a_1(\frac{p}{q}) + a_0 = 0$$

If we shift the constant term to the right hand side, factor a p and multiply by q^n, we get:

$$p(a_np^{n-1} + a_{n-1}qp^{n-2} + \ldots + a_1q^{n-1}) = -a_0q^n$$

We see that p times the integer quantity in parentheses equals $-a_0q^n$, so p divides a_0q^n. But p is coprime to q and therefore to q^n, so by (the generalized form of) **Euclid's lemma**, or first theorem, it must divide the remaining factor a0 of the product.

If we instead shift the leading term to the right hand side and multiply by qn, we get:

$$q(a_{n-1}p^{n+1} + a_{n-2}qp^{n-2} + \ldots + a_0q^{n-1} = -a_np^n$$

And for similar reasons, we can conclude that q divides a_n.

Example

For example, every rational solution of the equation

$$3x^3 - 5x^2 + 5x - 2 = 0$$

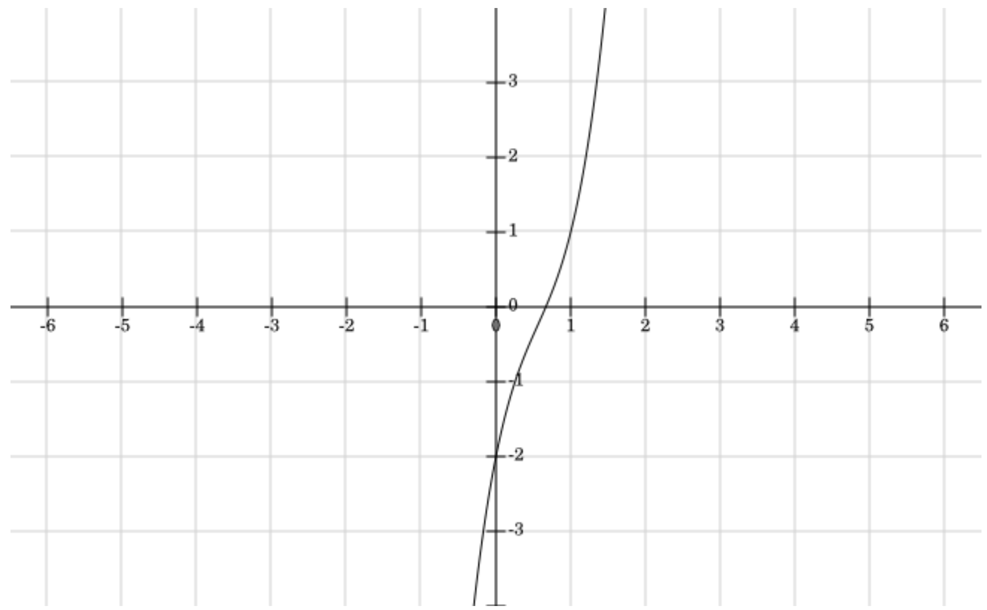
(shown graphically as: *Figure 4.17*) must be among the numbers symbolically indicated by:

$$\pm\frac{1,2}{1,3}$$

Which gives the list of possible answers:

$$1, -1, 2, -2, \frac{1}{3}, -\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}$$

Figure 4.17 Graph of 3x^3-5x^2+5x-2



One can also use the Rational Zeros Theorem to narrow down the candidates for solutions, then look to see which one is represented by the graphical form of the equation.

These root candidates can be tested using the Horner's method (for instance). In this particular case, there is exactly one rational root. If a root candidate does not satisfy the equation, it can be used to shorten the list of remaining candidates. For example, x = 1 does not satisfy the equation as the left hand side equals 1. This means that substituting x = 1 + t yields a polynomial in t with constant term 1, while the coefficient of t^3 remains the same as the coefficient of x^3. Applying the Rational Root Theorem thus yields the following possible roots for t:

$$t = \pm\frac{1}{1,3}$$

Therefore,

$$x = 1 + t = 2, 0, \frac{4}{3}, \frac{2}{3}$$

Root candidates that do not occur on both lists are ruled out. The list of rational root candidates has thus shrunk to just x = 2 and x = 2/3.

The Rule of Signs

The rule of signs gives an upper bound number of positive or negative roots of a polynomial.

KEY POINTS

- The rule of signs gives us an upper bound number of positive or negative roots of a polynomial. It is not a complete criterion, meaning that it does not tell the exact number of positive or negative roots.
- The rule states that if the terms of a polynomial with real coefficients are ordered by descending variable exponent, then the number of positive roots of the polynomial is either equal to the number of sign differences between consecutive nonzero coefficients, or is less by a multiple of 2.
- As a corollary of the rule, the number of negative roots is the number of sign changes after multiplying the coefficients of odd-power terms by −1 [f(-x)], or fewer than it by a multiple of 2.

The rule of **signs**, first described by René Descartes in his work La Géométrie, is a technique for determining the number of positive or negative real **root**s of a polynomial.

The rule gives us an upper bound number of positive or negative roots of a polynomial. However, it does not tell the exact number of positive or negative roots.

Positive Roots

In order to find the number of positive roots in a polynomial with only one variable, you must first arrange the polynomial by descending variable exponent. For example, x^2+x^3+x would have to be written x^3+x^2+x. Then, you must count the number of sign differences between consecutive nonzero coefficients. This number, and any number less than it by a multiple of 2, is your number of positive roots. It is important to note that multiple roots of the same value should be counted separately.

Negative Roots

Finding the negative roots is similar to finding the positive roots. The difference is, you must start by finding the coefficients of odd power (x^3 or x^5, not x^2 or x^4). Once you have located them, multiply each by -1. Then the procedure is the same, count the number of sign changes between consecutive nonzero coefficients. This number, and any number less than it by a multiple of 2, is your number of positive roots. Again it is important to note that multiple roots of the same value should be counted separately.

This can also be done by taking the function, f(x), and substituting the x for -x, f(-x). The reason we only bother to change the sign of the odd power coefficients, is because if we substitute in -x in an even power, it will just become a positive again. For example: $(-x)^3= (-x)(-x)(-x) = (-x)^3(-x)^2= (-x)(-x) = (x)^2$ The negative signs cancel out. By only multiplying the odd powered coefficients by -1, we are essentially saving ourselves a step.

Example

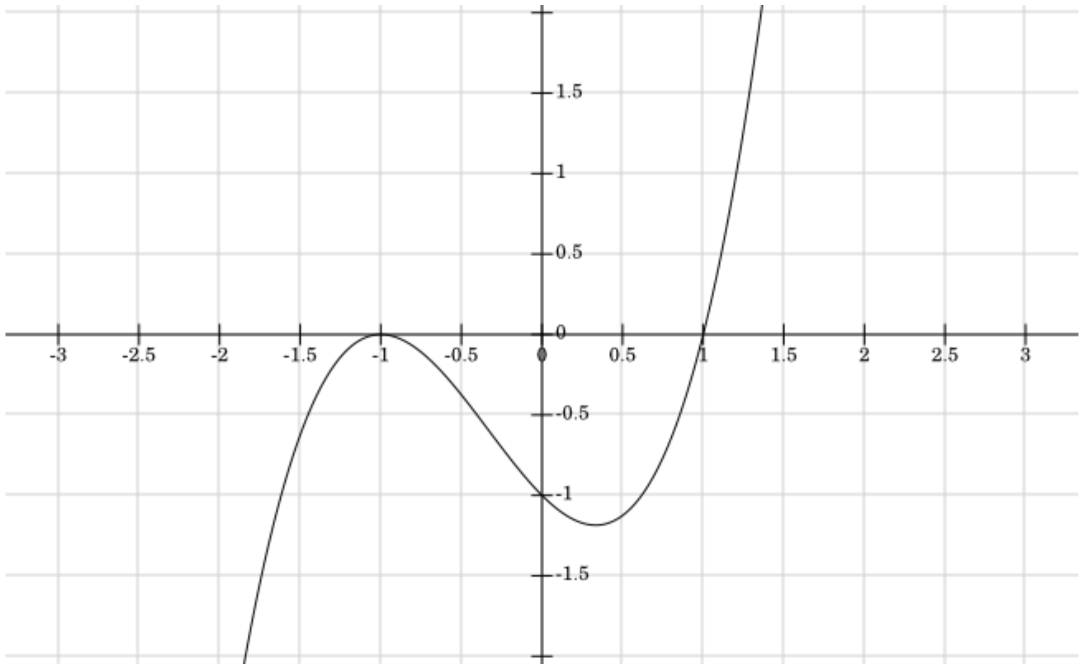
The polynomial:

$$f(x) = +x^3 + x^2 - x - 1$$

This function has one sign change between the second and third terms (the sequence of pairs of successive signs is ++, +−, −−). Therefore it has exactly one positive root. *Note: Don't forget that the first term HAS A SIGN, which, in this case, is positive. Be sure to remember to change its sign when looking for the negative roots. Next, we move on to finding the negative roots. To do this we apply what we learned from the rule of signs. Change the exponents of the odd powered coefficients. Once you have done this, you have obtained the second polynomial and are ready to find the number of negative roots. This second polynomial is shown below:

$$f(-x) = -x^3 + x^2 + x - 1$$

Figure 4.18 Plot of x^3+x^2-x-1



We can graphically see there are two solutions to this polynomial. This still fits with the rule of signs, as -1 is a negative root twice in the equation.

This polynomial has two sign changes (the sequence of pairs of successive signs is −+, ++, +−), so we know that it has at most two negative roots. From the rule of signs, we learned that the number or roots of either sign is the number of sign changes, or a multiple of two less than that. So this polynomial can have either 2 or 0 negative roots. We can validate this graphically, as shown in *Figure 4.18*.

So in this example the roots are 1, -1 and -1.

This can be checked by factoring the polynomial:

$f(x) = (x + 1)(x + 1)(x - 1)$, which simplifies into:

$f(x) = (x + 1)^2(x - 1)$

so the roots are −1,-1 and 1.

Complex Roots

A polynomial of the nth degree has exactly n roots. The minimum number of complex roots is equal to:

$n - (p + q)$

n = total number of roots in a polynomial; p = the maximum number of positive roots; q = the maximum number of negative roots.

A simple example is:

$x^2 + b$

To find the positive roots we count the sign changes. For this example, we will assume that b>0. This means that there are no positive roots. So p = 0. Now we look for negative roots. Since there are no odd powered coefficients, there are no changes to be made before looking for sign changes, therefore there are no negative roots. So, q = 0. Now we apply the complex root equation: n-(p+q): 2-(0+0) = 2. There are 2 complex roots.

Section 5

Rational Functions

https://www.boundless.com/algebra/polynomial-and-rational-functions/rational-functions/

Finding the Domain of a Rational Function

The domain of a rational function $f(x) = P(x)/Q(x)$ is the set of all points x for which the denominator Q(x) is not zero.

KEY POINTS

- A rational function is any function which can be written as the ratio of two polynomial functions.
- The domain of $f(x) = P(x)/Q(x)$ is the set of all points x for which the denominator Q(x) is not zero, where one assumes that the fraction is written in its lower degree terms, that is, P and Q have several factors of the positive degree.
- Domain restrictions can be determined by setting the denominator equal to zero and solving.

The Rational Function

A **rational function** is any function which can be written as the ratio of two polynomial functions. Neither the coefficients of the polynomials, nor the values taken by the function, are necessarily rational numbers.

In the case of one variable, x, a function is called a rational function if, and only if, it can be written in the form:

$f(x) = P(x)/Q(x)$

where P and Q are polynomial functions in x and Q is not the zero polynomial. The **domain** of f is the set of all points x for which the **denominator** Q(x) is not zero, where one assumes that the fraction is written in its lower degree terms; that is, P and Q have several factors of the positive degree.

Every polynomial function is a rational function with Q(x) = 1.

The Domain of a Rational Function

The denominator of the fraction cannot equal zero. Domain restrictions can be determined by setting the denominator equal to zero and solving. For example, the domain of $y = 1/x$ is comprised of all values of x for which x is not equal to zero.

The domain of the rational function $(x^2 - 3x - 2)/(x^2 - 4)$ includes all x not equal to +2 or -2 as can be seen in Figure 4.19.

An additional example is the domain of $(x + 3)/(x^2 + 2)$. The domain of this function is all real numbers, since for $x^2 + 2$ to equal 0, x^2 would need to equal -2, and this condition cannot be satisfied by a real number.

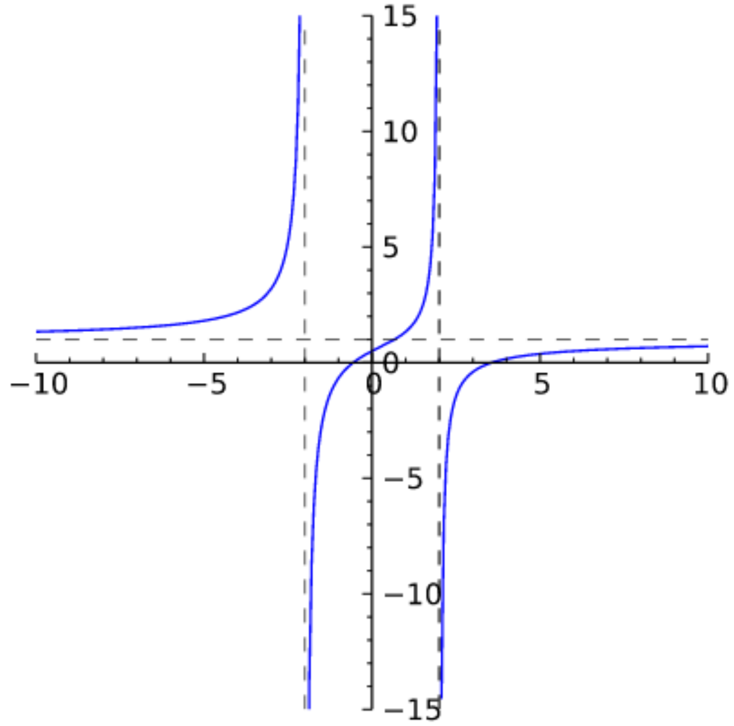


Figure 4.19 Example of a Rational Polynomial

The domain of this function includes all x not equal to +2 or -2.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/rational-functions/finding-the-domain-of-a-rational-function/

Asymptotes

An asymptote is a line such that the distance between a curve and the line approaches zero as both tend to infinity.

KEY POINTS

- An asymptote of a curve is a line such that the distance between the curve and the line approaches zero as they tend to infinity.
- There are potentially three kinds of asymptotes: horizontal, vertical and oblique asymptotes.
- A rational function has at most one horizontal asymptote or oblique (slant) asymptote, and possibly many vertical asymptotes.

In analytic geometry, an **asymptote** of a curve is a line such that the distance between the curve and the line approaches zero as they tend to infinity.

There are potentially three kinds of asymptotes: horizontal, vertical and **oblique** asymptotes. For curves given by the graph of a function $y = f(x)$, horizontal asymptotes are horizontal lines that the graph of the function approaches as x tends to $+\infty$ or $-\infty$. Vertical

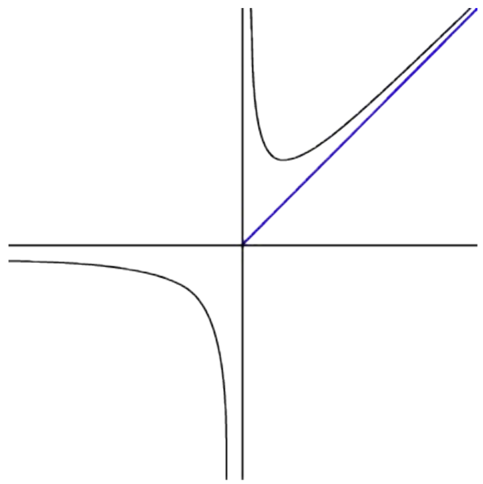
asymptotes are vertical lines near which the function grows without bound.

A Simple Example

Consider the graph of the equation y = 1/x shown to the right. The coordinates of the points on the curve are of the form (x, 1/x) where x is a number other than 0. For example, the graph contains the points (1, 1), (2, 0.5), (5, 0.2), (10, 0.1),... As the values of x become larger and larger, say 100, 1000, 10,000..., putting them far to the right of the illustration, the corresponding values of y,. 01,.001,.0001,..., become infinitesimal relative to the scale shown. But no matter how large x becomes, its reciprocal 1/x is never 0, so the curve never actually touches the x-axis. Similarly, as the values of x become smaller and smaller, say.01,.001,.0001,..., making them infinitesimal relative to the scale shown, the corresponding values of y, 100, 1000, 10,000..., become larger and larger. So the curve extends farther and farther upward as it comes closer and closer to the y-axis. Thus, both the x and y-axes are asymptotes of the curve. These ideas are part of the basis of concept of a limit in mathematics.

Asymptotes of functions

Figure 4.20 Asymptotes



The graph of a function with a horizontal, vertical, and oblique asymptote.

Asymptotes can be classified into horizontal, vertical and oblique asymptotes depending on orientation. Examples are shown in *Figure 4.20*.

Horizontal asymptotes are horizontal lines that the graph of the function approaches as x tends to +∞ or −∞. As the name indicate they are parallel to the x-axis. Functions may lack horizontal asymptotes on either or both sides, or may have one horizontal asymptote that is the same in both directions.

Vertical asymptotes are vertical lines (perpendicular to the x-axis) near which the function grows without bound. A common example of a vertical asymptote is the case of a **rational function** at a point x such that the denominator is zero and the numerator is non-zero.

Oblique asymptotes are diagonal lines so that the difference between the curve and the line approaches 0 as x tends to +∞ or −∞. When a linear asymptote is not parallel to the x- or y-axis, it is called an oblique asymptote or slant asymptote.

Asymptotes of many elementary functions can be found without the explicit use of limits (although the derivations of such methods typically use limits).

Asymptotes of Rational Functions

A rational function has at most one horizontal asymptote or oblique (slant) asymptote, and possibly many vertical asymptotes. The degree of the numerator and degree of the denominator determine whether or not there are any horizontal or oblique asymptotes. When the numerator of a rational function has degree exactly one greater than the denominator, the function has an oblique (slant) asymptote. The asymptote is the polynomial term after dividing the numerator and denominator. This phenomenon occurs because when dividing the fraction, there will be a linear term, and a remainder. The vertical asymptotes occur only when the denominator is zero. If both the numerator and denominator are zero, the multiplicities of the zero are compared.

Solving Problems with Rational Functions

In the case of rational functions, the x-intercepts exist when the numerator is equal to 0.

KEY POINTS

- A rational function is defined as the ratio of two real polynomials with the condition that the polynomial in the denominator is not a zero polynomial.
- The x-intercepts, also known as zeros of the function or real roots, can be more than one x-intercept. On graphs, x-intercepts are points where a graph intersects the x-axis. Thus, x-intercepts are x-values for which the function has a value of zero.
- In the case of rational functions, the x-intercepts exist when the numerator is equal to 0. In the case of rational functions, the x-intercepts exist when the numerator is equal to 0. For $f(x) = P(x)/Q(x)$, if $P(x) = 0$, then $f(x) = 0$.

A **rational function** is defined in similar fashion as a rational number is defined in terms of **numerator** and **denominator**. It is defined as the ratio of two real polynomials with the condition that the polynomial in the denominator is not a zero polynomial.

$f(x) = P(x)/Q(x)$

An example of a rational functions is: $f(x) = (x + 1)/(2x^2 - x - 1)$

The x-intercepts are also known as zeros of the function or real roots. There can be more than one x-intercept. On graphs, x-intercepts are points where a graph intersects the x-axis. Thus, x-intercepts are x-values for which the function has a value of zero:

$f(x) = 0$

In the case of rational functions, the x-intercepts exist when the numerator is equal to 0. For $f(x) = P(x)/Q(x)$, if $P(x) = 0$, then $f(x) = 0$.

Example 1

Find the x-intercepts of the function f(x) = 1/x. Here, the numerator is 1 and cannot be zero. Thus, this function does not have any x-intercepts.

Example 2

Find the x-intercepts of the function $(x^2 - 3x + 2)/(x^2 - 2x - 3)$. The numerator of this rational function can be factored: $x^2 - 3x + 2 = (x - 1)(x - 2)$. This polynomial equals 0 when x = 1 or 2. The x-intercepts can thus be found at 1 and 2.

Example 3

Find the roots of $(x^3 - 2x)/(2x^2 - 10)$. Equating the numerator to 0, it can be seen that the roots of this function are 0 and the positive and negative square roots of 2. These values can also be identified from *Figure 4.21*.

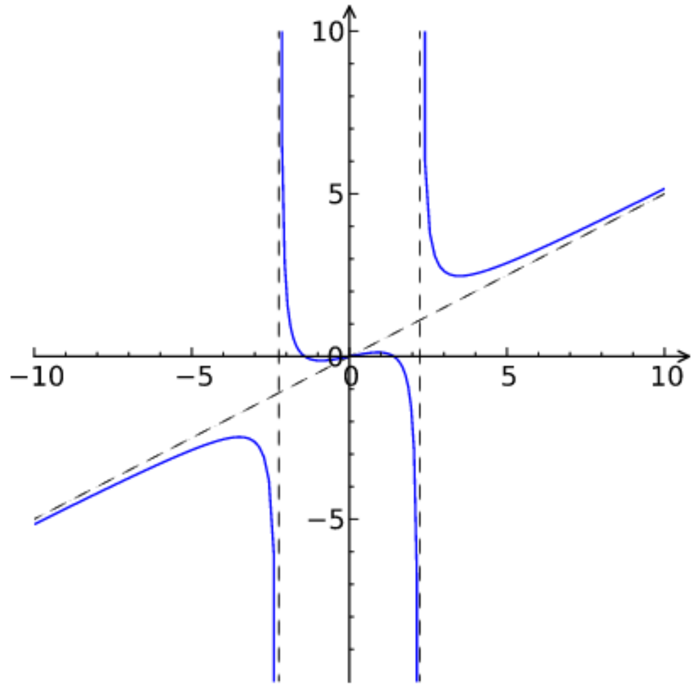


Figure 4.21
Example of a Rational Function
This function has three x-intercepts.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/rational-functions/solving-problems-with-rational-functions/

Section 6

Inequalities

Polynomial Inequalities

Rational Inequalities

https://www.boundless.com/algebra/polynomial-and-rational-functions/inequalities/

Polynomial Inequalities

Polynomials can be expressed as inequalities, the solutions for which can be determined from the polynomial's zeros.

KEY POINTS

- To solve a polynomial inequality, first rewrite the polynomial in factored form to find its zeros.
- For each zero, input the value of the zero in place of x in the polynomial. Determine the sign (positive or negative) of the polynomial as it passes the zero in the rightward direction.
- Determine the intervals between these roots which satisfy the inequality.

Like any other function, a polynomial may be written as an **inequality**, giving a large range of solutions.

The best way to solve a polynomial inequality is to find its zeros. At these points, the polynomial's value goes from negative to positive or positive to negative. The easiest way to find the zeros of a polynomial is to express it in factored form.

Consider the polynomial inequality:

$$x^3 + 2x^2 - 5x - 6 > 0$$

This can be expressed as the product of these three terms:

$$(x-2)(x+1)(x+3) > 0$$

The three terms reveal zeros at x=-3, x=-1, and x=2. We know that the lower limit of the inequality crosses the x axis at each of these x values, but now have to determine which direction (positive or negative) it takes at each crossing.

x+3>0 for x>-3

x+1>0 for x>-1

x-2>0 for x>2

Thus, as the polynomial crosses the x axis at x=-3, the term (x+3) equals 0, becoming positive to the right. At the same time, (x+1) and (x-2) are negative. The product of a positive and two negatives is positive, so we can conclude that the polynomial becomes positive as it passes x=-3.

The next zero is at x=-1. From the paragraph above, we know that the polynomial is positive as it approaches its next zero, but we can use the same reasoning for proof. At x=-1, (x+1) equals 0, becoming positive to the right. The term (x+3) is positive, while (x-2) is negative. The product of two positives and a negative is negative, so we can conclude that the polynomial becomes negative as it passes x=-1.

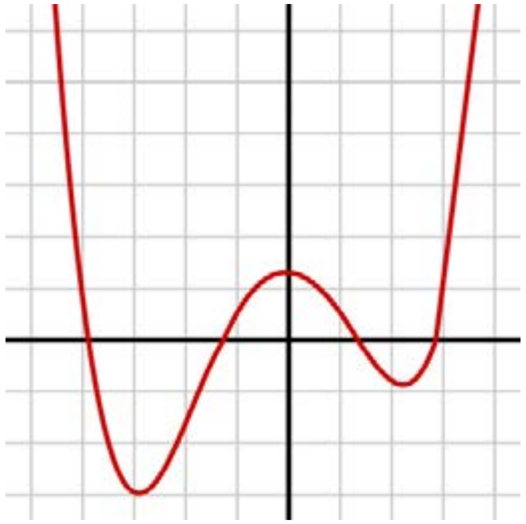
The same process can be used to show that the polynomial becomes positive again at x=2.

Recalling the initial inequality, we can now determine the solution of exactly where the polynomial is greater than 0. Because there is no zero to the left of x=-3, we can assume that the polynomial is negative for all x values -∞ to -3. The polynomial is positive from x=-3 to x=-1 before becoming negative once more. It becomes positive at x=2, and because there are no more zeros to the right, we can assume the polynomial remains positive as x approaches ∞ (*Figure 4.22*).

Thus, the solution is: (-3, -1), (2,∞)

For inequalities that are not expressed relative to zero, expressions can be added or subtracted from each side to take it into the desired form.

Figure 4.22 A fourth-degree polynomial



This polynomial has four roots. It is positive in three segments and negative in two. If it were a polynomial inequality with the condition that all values are greater than 0, the two negative segments would be removed.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/inequalities/polynomial-inequalities/

Rational Inequalities

Rational inequalities can be solved much like polynomial inequalities.

KEY POINTS

- First factor the numerator and denominator polynomial to reveal the zeros in each.
- Substitute x with a zero (root) to determine whether the rational function is positive or negative to the right of that point. Repeat for all zeros.
- The intervals that satisfy the inequality symbol will be the answer. Note that for any ≥ or ≤, the interval will only be closed to include the zero if the zero is found in the numerator. If the zero is found in the denominator, that point is undefined, and cannot be included in the solution.

As with solving polynomial inequalities, the first step to solving rational **inequalities** is to find the **zeros**.

Because a rational expression consists of the ratio of two polynomials, the zeroes will be needed for both.

The zeros in the numerator are x values at which the rational inequality crosses from negative to positive or from positive to negative. The zeros in the denominator are x values are at which the rational inequality is undefined, the result of dividing by 0 (*Figure 4.23*).

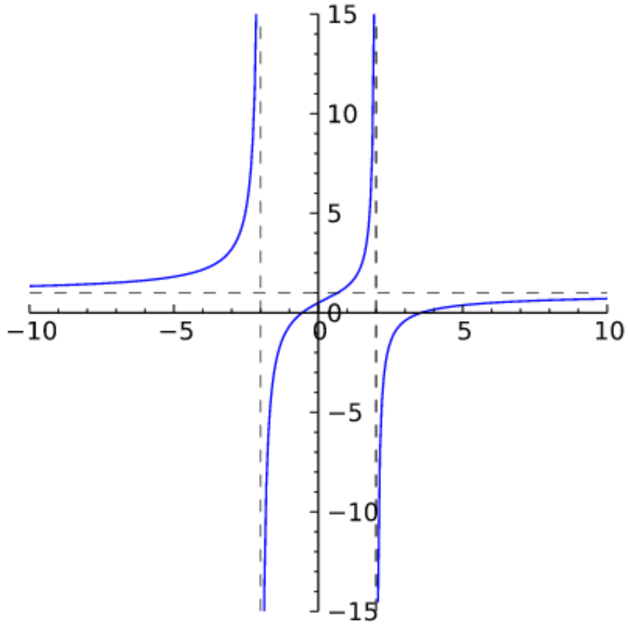


Figure 4.23
Example of a rational polynomial
For x values that are zeros for the numerator polynomial, the rational function overall is equal to zero. For x values that are zeros for the denominator polynomial, the rational function is undefined, with a vertical asymptote forming instead.

Consider the rational inequality:

$$\frac{x^2+2x-3}{x^2-4} > 0$$

This equation can be factored to give:

$$\frac{(x+3)(x-1)}{(x+2)(x-2)} \geq 0$$

The numerator has zeros at x=-3 and x=1. The denominator has zeros at x=-2 and x=2.

As x crosses rightward past -3, (x+3) becomes positive. At that same point, (x-1), (x+2), and (x-2) are all negative. The product of a positive and three negatives is negative, so the rational expression becomes negative as it crosses x=-3 in the rightward direction.

The same process can be used to determine that the rational expression is positive after passing the zero at x=-2, is negative after passing x=1, and is positive after passing x=2.

Thus we can conclude that for x values on the open interval from -∞ to -3, the rational expression is negative. From -3 to -2, it is positive; from -2 to 1 it is negative; from 1 to 2 it is positive, and from 2 to ∞ it is negative.

Because the inequality is written as ≥ 0 as opposed to > 0, we will need to evaluate the x values at zeros to determine whether the function is defined.

In the case of x=-2 and x=2, the rational function has a denominator equal to 0 and becomes undefined.

In the case of x=-3 and x=1, the rational function has a numerator equal to 0, which makes the function overall equal to 0, making it inclusive in the solution.

Thus, the full solution is:

$[-3,-2),[1,2)$

Section 7

Variation and Problem Solving

Direct Variation

Inverse Variation

Combined Variation

https://www.boundless.com/algebra/polynomial-and-rational-functions/variation-and-problem-solving/

Direct Variation

When two variables change proportionally, or are directly proportional, to each other, they are said to be in direct variation.

KEY POINTS

- The ratio of variables in direct variation is always constant
- Direct variation between variables is easily modeled using a linear graph.
- The equation relating directly varying variables to a constant can be rearranged to slope-intercept form.

When two variables change proportionally to each other, they are said to be in direct variation. This can also be called directly proportional.

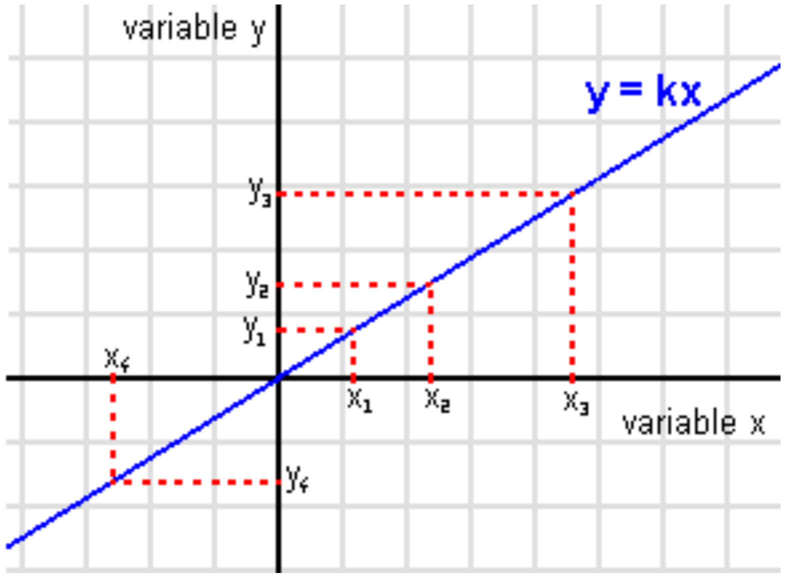
For example, a toothbrush costs $2. Purchasing five toothbrushes would cost $10; purchasing 10 toothbrushes would cost $20. No matter how many toothbrushes purchased, the ratio will always remain: $2 per toothbrush. Thus we can say that cost varies directly as the value of toothbrushes.

Direct variation is easily illustrated using a linear graph. Knowing that the relationship between two variables is constant, we can show their relationship as (*Figure 4.24*):

$$\frac{y}{x} = k$$

where k is a constant known as the constant of proportionality.

Figure 4.24 Direct Variation



The line y=kx is an example of direct variation between variables x and y. For all points on the line, y/x=k.

This can be rearranged to slope-intercept format:

$$y = kx$$

In this case, the y-intercept is equal to 0.

Revisiting the example with toothbrushes and dollars, we can define the x axis as number of toothbrushes and the y axis as number of dollars. Doing so, the variables would abide by the relationship:

$$\frac{y}{x} = 2$$

Any augmentation of one variable would lead to an equal augmentation of the other. For example, doubling y would result in the doubling of x.

Source: https://www.boundless.com/algebra/polynomial-and-rational-functions/variation-and-problem-solving/direct-variation/

Inverse Variation

Indirect variation is used to describe the relationship between two variables when their product is constant.

KEY POINTS

- The ratio of variables in direct variation is always constant.
- Direct variation between variables is depicted by an hyperbola.
- The equation relating indirectly varying variables to a constant can be rearranged to hyberbolic form.

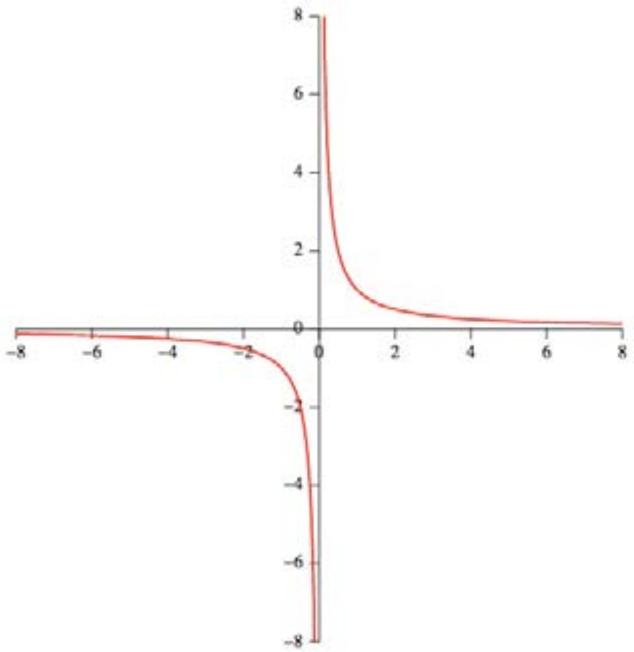
Inverse variation is the opposite of direct variation. In the case of inverse variation, the increase of one variable leads to the decrease of another.

Consider a car driving on a flat surface at a certain speed. If the driver shifts into neutral gear, the car's speed will decrease at a **constant** rate as time increases, eventually coming to a stop.

Inverse variation can be illustrated, forming a graph in the shape of a **hyperbola** (*Figure 4.25*). Knowing that the relationship between the two variables is constant, we can show that their relationship is:

$yx = k$

Figure 4.25 Indirect variation



This hyperbola shows the indirect variation of variables x and y.

where k is a constant known as the constant of proportionality. Note that as long as k is not equal to 0, neither x nor y can ever equal 0 either.

We can rearrange the above equation to place the variables on opposite sides:

$y = k/x$

Revisiting the example of the decelerating car, let's say it starts at 50 miles per hour and slows at a constant rate. If we define y as its speed in miles per hour, and x as time, the relationship between x and y can be expressed as:

$y = 50/x$

Note that realistically, other factors (e.g., friction), will influence the rate of deceleration. Other constants can be incorporated into the equation for the sake of accuracy, but the overall form will remain the same.

Combined Variation

Combined variation describes the relationship between three or more variables that vary directly and inversely with one another.

KEY POINTS

- There must be a minimum of three related variables for their relationship to be one of combined variation.
- Among the three or more related variables, one must directly vary with another and inversely vary with a third in order for the relationship to be one of combined variation.
- An example of combined variation in the physical world is the Combined Gas Law, which relates pressure, temperature, volume, and moles (amount of molecules) of a gas.

Combined variation is used to describe the relationship between three or more variables that vary directly and inversely with one another. Before go deeper into the concept of combined variation, it is important to first understand what direct and inverse variation mean.

Direct and Inverse Variation

Simply put, two variables are in direct variation when the same thing that happens to one variable happens to the other. If x and y are in direct variation, and x is doubled, then y would also be doubled. The two variables may be considered **directly proportional**.

Two variables are said to be in inverse variation, or are inversely proportional, when an operation of change is performed on one variable and the opposite happens to the other. For example, if x and y are inversely proportional, if x is doubled, then y is halved.

Combined Variation

To have variables that are in combined variation, the equation must have variables that are in both direct and inverse variation, as shown in the example below.

Consider the equation:

$$z = k(\frac{x}{y})$$

where x, y, and z are variables and k is a **constant** known as the proportionality constant.

In this example, z varies directly as x and inversely as y.

Given values for any three of x, y, z, and k, the fourth can be found by substitution. For example, if z=12, x=4 and y=2, we can solve for k:

$$12 = k\frac{4}{2}$$

$$k = 6$$

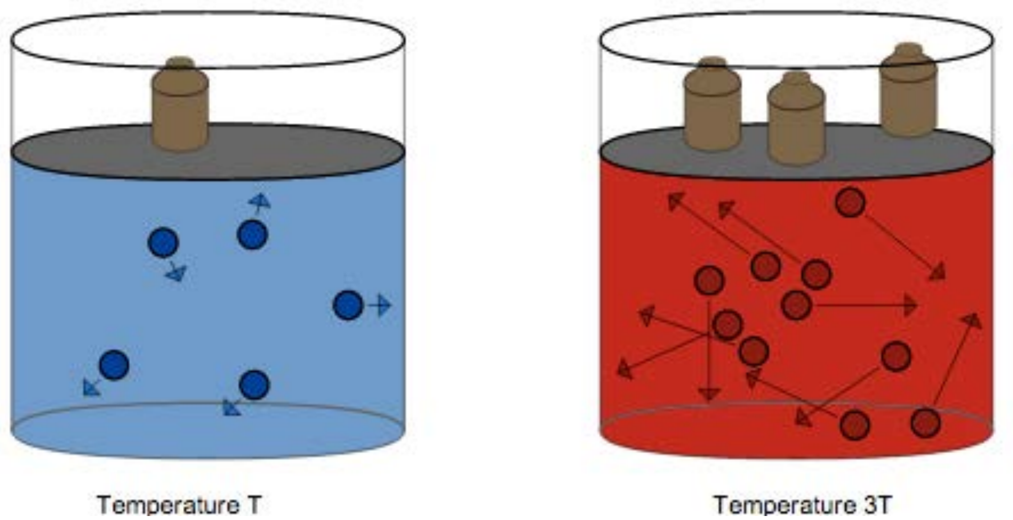
Practical Application

A practical example of combined variation is the Combined Gas Law, which relates the pressure (p), volume (v), moles (n), and temperature (T) of a sample of gas:

$$PV = nRT$$

where R is a constant (*Figure 4.26*).

Figure 4.26 Illustration of Gay-Lussac's Law, derived from the Combined Gas Law



A constant amount of gas will exert pressure that varies directly with temperature. In this illustration, volume is held constant by an increased mass weighing down the lid of the container. If not for that extra mass, the lid would raise, increasing the volume and relieving the pressure.

Solving for P, we can determine the variation of the variables.

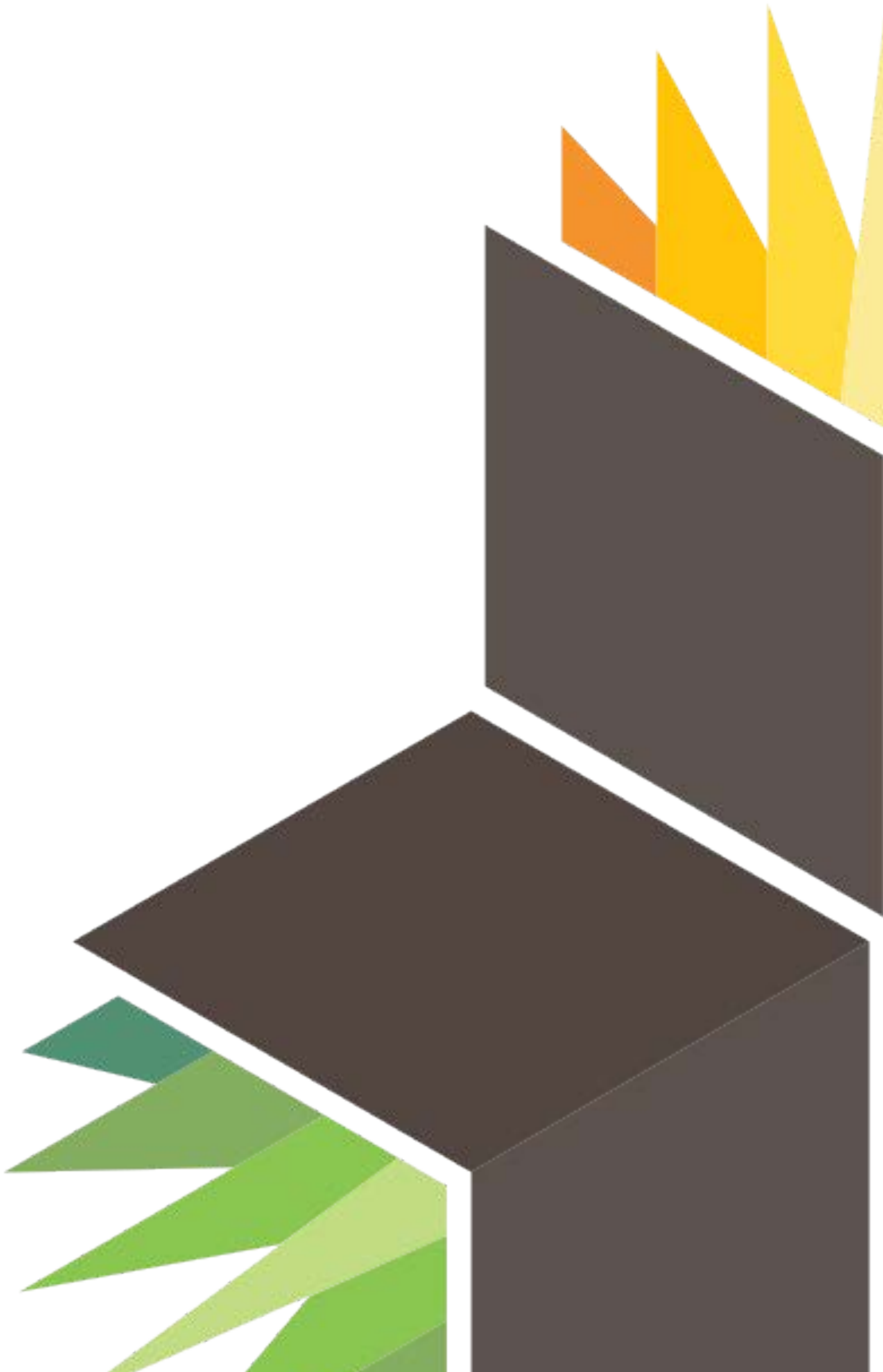
$$P = \frac{nRT}{V}$$

In the above equation, P varies directly with n and T, and inversely with V. Thus, pressure increases as temperature and moles increase. What's more, pressure decreases as volume increases.

Chapter 5

Exponents and Logarithms

https://www.boundless.com/algebra/exponents-and-logarithms/



Section 1

Inverse Functions

https://www.boundless.com/algebra/exponents-and-logarithms/inverse-functions/

Inverses

Logarithm reverses exponentiation. The complex logarithm is the inverse function of the exponential function applied to complex numbers.

KEY POINTS

- In trigonometric functions, a positive exponent applied to the function's abbreviation means that the result is raised to that power, while an exponent of −1 denotes the inverse function.
- An inverse function is a function that undoes another function: If an input x into the function f produces an output y, then putting y into the inverse function g produces the output x, and vice versa (i.e., f(x)=y, and g(y)=x).
- The logarithm to base b is the inverse function of f(x) = bx: $\log_b (b)^x = \quad x \log_b(b) = \quad x.$
- The natural logarithm ln(x) is the inverse of the exponential function ex: $b = e^{lnb}$.

In mathematics, an **inverse function** is a function that undoes another function: If an input x into the function f produces an output y, then putting y into the inverse function g produces the output x, and vice versa (i.e., f(x)=y, and g(y)=x). More directly, g(f(x))=x means that g(x) composed with f(x) leaves x unchanged. A function f that has an inverse is called invertible; the inverse function is then uniquely determined by f and is denoted by f^{-1} (read f inverse, not to be confused with exponentiation).

Exponential Notation

Placing an integer superscript after the name or symbol of a function, as if the function were being raised to a power, commonly refers to repeated function composition rather than repeated multiplication. Thus $f^3(x)$ may mean f(f(f(x))); in particular, $f^{-1}(x)$ usually denotes the inverse function of f. Iterated functions are of interest in the study of fractals and dynamical systems. A special syntax applies to the trigonometric functions: a positive exponent applied to the function's abbreviation means that the result is raised to that power, while an exponent of −1 denotes the inverse function. That is, $\sin^2 x$ is just a shorthand way to write $(\sin x)^2$ without using parentheses, whereas $\sin^{-1} x$ refers to the inverse function of the sine, also called arcsin x. There is no need for a shorthand for the reciprocals of trigonometric functions since each has its own name and abbreviation; for example, $1/(\sin x) = (\sin x)^{-1} = \csc x$. A similar convention applies to logarithms, where $\log^2 x$ usually means $(\log x)^2$, not log log x.

Exponents and Logarithms

The formula for the logarithm of a power says in particular that for any number x:

$$log_b(b)^x = \quad x \log_b(b) = \quad x$$

In prose, taking the x-th power of b and then the base-b logarithm gives back x. Conversely, given a positive number y, the formula says that first taking the logarithm and then exponentiating gives back y. Thus, the two possible ways of combining (or composing) logarithms and exponentiation give back the original number. Therefore, the logarithm to base b is the inverse function of f(x) = bx. Inverse functions are closely related to the original functions.

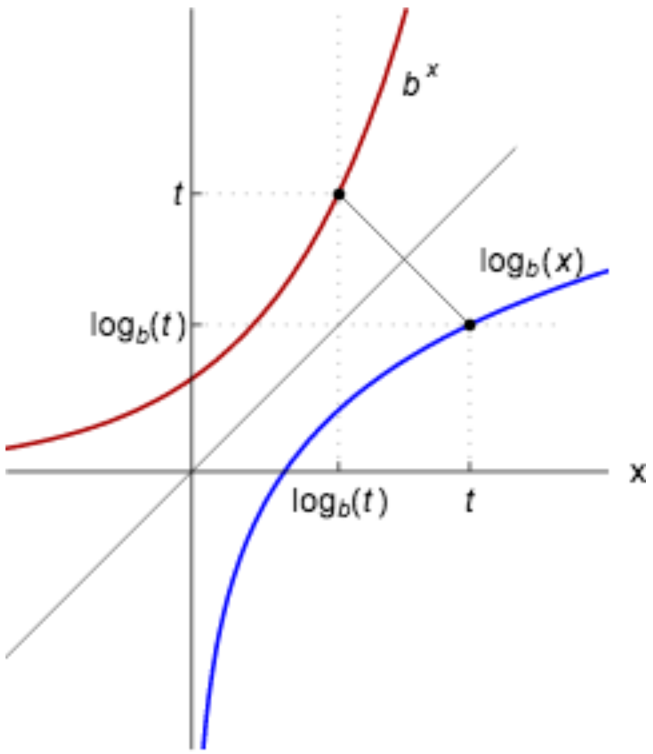


Figure 5.1
Logarithm Function
The graph of the logarithm function logb(x) (blue) is obtained by reflecting the graph of the function bx (red) at the diagonal line (x = y).

Their graphs correspond to each other upon exchanging the x- and the y-coordinates (or upon reflection at the diagonal line x = y), as shown here ([Figure 5.1]()). A point (t, u = b^t) on the graph of f yields a point (u, t = $\log_b u$) on the graph of the logarithm and vice versa. As a consequence, $\log_b(x)$ diverges to infinity (gets bigger than any given number) if x grows to infinity, provided that b is greater than one. In that case, $\log_b(x)$ is an increasing function. For b < 1, $\log_b(x)$ tends to minus infinity instead. When x approaches zero, $\log_b(x)$ goes to minus infinity for b > 1 (plus infinity for b < 1, respectively).

In the same way as the logarithm reverses exponentiation, the complex logarithm is the inverse function of the exponential function applied to complex numbers.

The natural logarithm ln(x) is the inverse of the exponential function e^x. It is defined for b > 0, and satisfies:

$$b = e^{lnb}$$

If b^x is to preserve the logarithm and exponent rules, then one must have:

$$b^x = (e^{lnb})^x = e^{x \cdot lnb}$$

for each real number x. This can be used as an alternative definition of the real number power b^x.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/inverse-functions/inverses/

One-to-One Functions

A one-to-one function, also called an injective function, never maps distinct elements of its domain to the same element of its codomain.

KEY POINTS

- A one-to-one function has an inverse function, if a function is not one-to-one it cannot have an inverse function as it yields multiple outputs.
- Domain restriction can allow a function to become one-to-one, as in the case of $f(x) = x^2$.
- An easy way to check if a function is a one-to-one function is by graphing it and then performing the horizontal line test.

A one-to-one function, also called an **injective function**, never maps distinct elements of its domain to the same element of its co-domain. In other words, every element of the function's co-domain is mapped to by at most one element of its domain. If, in addition, all of the elements in the co-domain are in fact mapped to by some element of the domain, then the function is said to be bi-jective. An injective function is also said to be a one-to-one function.

Occasionally, an injective function from X to Y is denoted f: X ↣ Y, using an arrow with a barbed tail. The set of injective functions from X to Y may be denoted Y^X using a notation derived from that

used for falling factorial powers, since if X and Y are finite sets with respectively m and n elements, the number of injections from X to Y is nm. A function f that is not injective is sometimes called many-to-one. However, this terminology is also sometimes used to mean "single-valued", i.e., each argument is mapped to at most one value. This is the case for any function, but is used to stress the opposition with multi-valued functions, which are not true functions.

An easy way to check if a function is a one-to-one is by graphing it and then performing the horizontal line test. If any horizontal line intersects the graph in more than one point, the function is not injective. To see this, note that the points of intersection have the same y-value, because they lie on the line, but different x values, which by definition means the function cannot be injective.

The exponential function $\exp : \mathbb{R} \to \mathbb{R}$ defined by $\exp(x) = e^x$ is injective, but not surjective as no real value maps to a negative number. The natural logarithm function $\ln : (0, \infty) \to \mathbb{R}$ defined by $x \mapsto \ln x$ is injective.

The exponential function, $f(x) = x^2$, without any domain restriction, is not one-to-one. It forms a parabola and fails the horizontal line test ([Figure 5.3](#)). Since it is not a one-to-one function, it cannot have an inverse. Looking at another exponential function: $f(x) = 10^x$

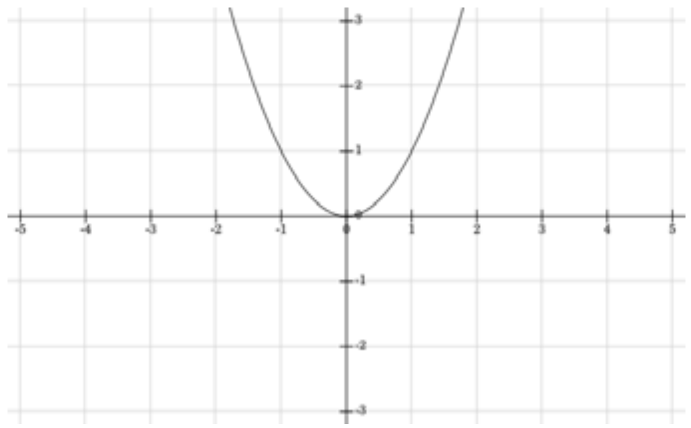


Figure 5.3 Graph of f(x)=x^2

This function forms a parabola and therefore fails the the horizontal line test.

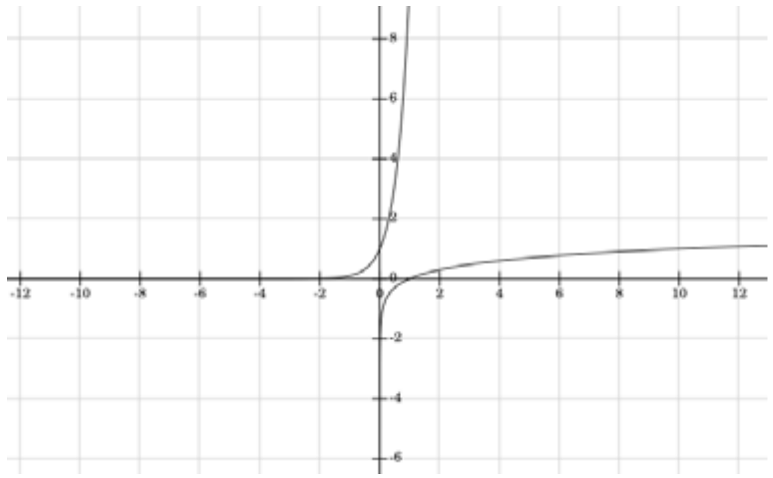


Figure 5.2 f(x)=10^x and its inverse f(x)=10^x

f(x)=10^x is a one-to-one function and is shown here with its inverse, f(x)=10^x.

([Figure 5.2](#)), it is a one-to-one function and therefore, has an inverse which is: $f^{-1}(x) = log(x)$.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/inverse-functions/one-to-one-functions/

Finding Formulas for Inverses

To find the inverse function, switch the x and y values, and then solve for y.

KEY POINTS

- An inverse function reverses the inputs and outputs.
- To find the inverse formula of a function, write it in the form of y and x, switch y and x, and then solve for y.
- Some functions have no inverse function, as a function cannot have multiple outputs.

Inverse function, $f^{-1}(x)$, is defined as the inverse function of f(x) if it consistently reverses the f(x) process. That is, if f(x) turns a into b, then $f^{-1}(x)$ must turn b into a. More concisely and formally, $f^{-1}(x)$ is the inverse function of f(x) if $f(f^{-1}(x)) = x$.

Finding an Inverse Function

In general, given a function, how do you find its inverse function? Remember that an inverse function reverses the inputs and outputs. When we graph functions (*Figure 5.4*), we always represent the incoming number as x and the outgoing number as y. So to find the inverse function, switch the x and y values, and then solve for y.

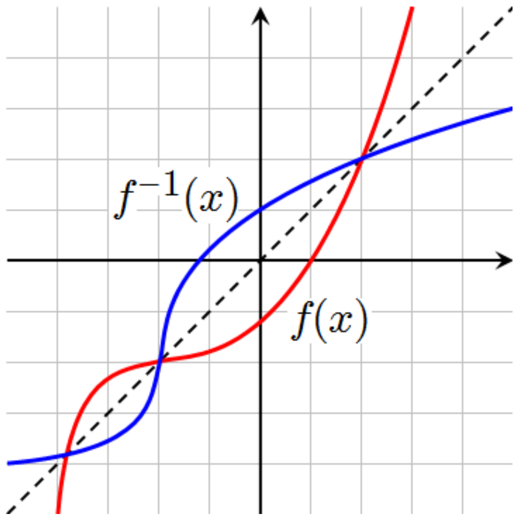


Figure 5.4 Graph of the Inverse

The graphs of y = f(x) and y = f−1(x). The dotted line is y = x.

Example

Building and Testing an Inverse Function

Find the inverse function of : $f(x) = 2^x$

a.: Write the function as: $y = 2^x$

b.: Switch the x and y variables: $x = 2^y$

c.: Solve for y: $log_2 x = log_2 2^y$

$log_2 x = y log_2 2$

$log_2 x = y$

So: $f^{-1}(x) = log_2(x)$

Test to make sure this solution fills the definition of an inverse function.

a.: Pick a number, and plug it into the original function. 2→f(x)→4.

b.: See if the inverse function reverses this process. 4→f−1(x)→2. ✓

Some functions have no inverse function because of the rule of consistency. For instance, consider the function y=x2. This function takes both 3 and −3 and turns them into 9. No problem: a function is allowed to turn different inputs into the same output. However, what does that say about the inverse of this particular function? In order to fulfill the requirement of an inverse function, it would have to take 9, and turn it into both 3 and −3, which is the one and only thing that functions are not allowed to do. Hence, the inverse of this function would not be a function at all!

Source: https://www.boundless.com/algebra/exponents-and-logarithms/inverse-functions/finding-formulas-for-inverses/

Composition and Composite Functions

A composite function represents, in one function, the results of an entire chain of dependent functions.

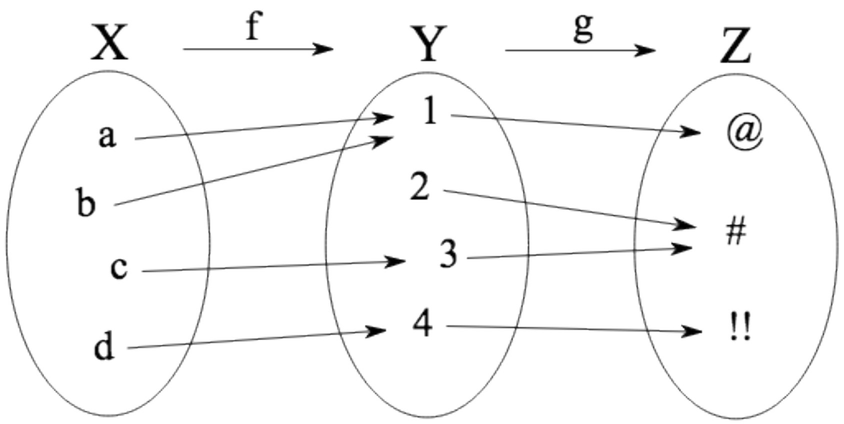
KEY POINTS

- The composition of functions is always associative. That is, if f, g, and h are three functions with suitably chosen domains and codomains, then f ∘ (g ∘ h) = (f ∘ g) ∘ h, where the parentheses serve to indicate that composition is to be performed first for the parenthesized functions.
- Functions can be inverted and then composed, giving the notation of: $(f' \circ g')(x)$.
- Functions can be composed and then inverted, yielding the following notation: $(f \circ g)'(x)$.

In mathematics, **function composition** is the application of one function to the results of another *Figure 5.5*. For instance, the functions f: X → Y and g: Y → Z can be composed by computing the output of g when it has an argument of f(x) instead of x.

Intuitively, if z is a function g of y and y is a function f of x, then z is a function of x. Thus one obtains a composite function g ∘ f: X → Z defined by (g ∘ f)(x) = g(f(x)) for all x in X. The notation g ∘ f is read

Figure 5.5 Composition of Functions



g ∘ f, the composition of f and g. For example, $(g \circ f)(c) = \#$.

as "g circle f", or "g composed with f", "g after f", "g following f", or just "g of f". The composition of functions is always associative. That is, if f, g, and h are three functions with suitably chosen domains and codomains, then $f \circ (g \circ h) = (f \circ g) \circ h$, where the parentheses serve to indicate that composition is to be performed first for the parenthesized functions. Since there is no distinction between the choices of placement of parentheses, they may be safely left off.

The functions g and f are said to commute with each other if $g \circ f = f \circ g$. In general, composition of functions will not be commutative. Commutativity is a special property, attained only by particular functions, and often in special circumstances.

Let's look at it in a basic sense:

You are working in the school cafeteria, making peanut butter sandwiches for today's lunch. The more classes the school has, the more children there are. The more children there are, the more sandwiches you have to make. The more sandwiches you have to make, the more pounds (lbs) of peanut butter you will use. The more peanut butter you use, the more money you need to budget for peanut butter...and so on. Each sentence in this little story is a function. Mathematically, if c is the number of classes and h is the number of children, then the first sentence asserts the existence of a function h(c).The principal walks up to you at the beginning of the year and says "We're considering expanding the school. If we expand to 70 classes, how much money do we need to budget? What if we expand to 75? How about 80?" For each of these numbers, you have to calculate each number from the previous one, until you find the final budget number.

But going through this process each time is tedious. What you want is one function that puts the entire chain together: "You tell me the number of classes, and I will tell you the budget." This is a composite function—a function that represents in one function the results of an entire chain of dependent functions. Since such chains are very common in real life, finding composite functions is a very important skill.

Inverses and composition

If f is an invertible function with domain X and range Y, then

$f^{-1}(f(x)) = x$ for every $x \quad \epsilon \quad X$

This statement is equivalent to the first of the above-given definitions of the inverse, and it becomes equivalent to the second definition if Y coincides with the codomain of f.

Let's go through the relationship between inverses and composition in this example, let's take two functions, compose and invert them, find: $(f \circ g)'(x)$

$f(x) = 2^x$

$g(x) = x + 1$

First, compose them:

$(f \circ g)(x) = f(x + 1)$

Then, invert it:

$x = 2^{(y+1)}$

$log_2(x) = log_2(2^{(y+1)})$

$log_2(x) = (y + 1)log_2(2)$

$log_2(x) = (y + 1)1$

$y = log_2(x) - 1$

$(f \circ g)'(x) = log_2(x) - 1$

As you can see here, a composite function with logs and exponents can also be inverted using the tools we've already learned.

Restricting Domains

Domain restriction is important for inverse functions of exponents and logarithms because sometimes we need to find an unique inverse.

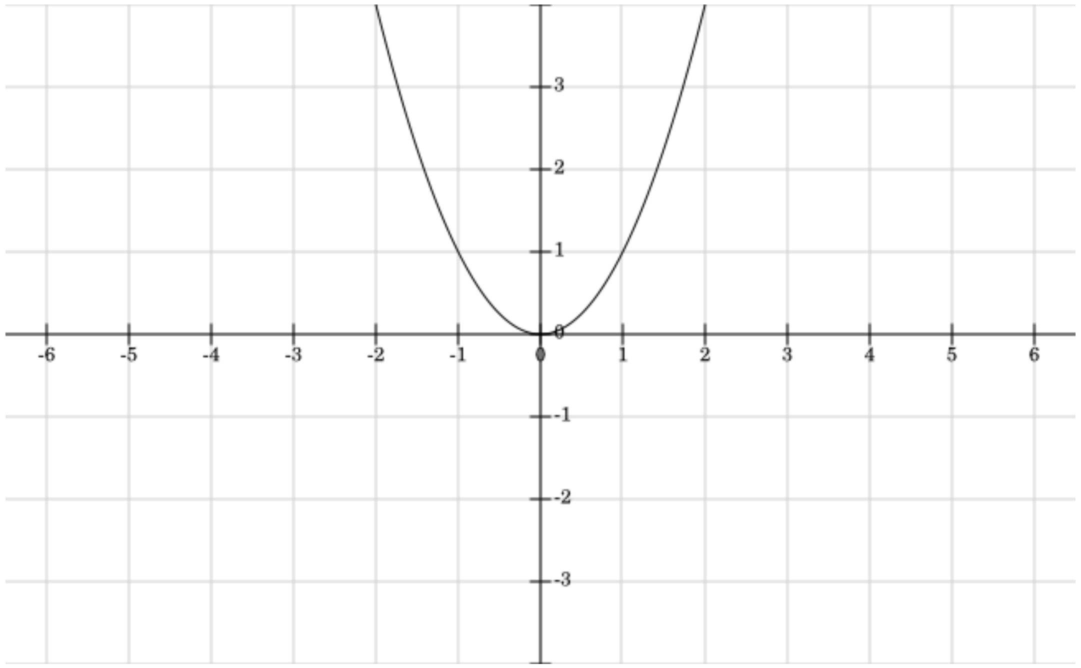
KEY POINTS

- $f^{-1}(x)$ is defined as the inverse function of f(x) if it consistently reverses the f(x) process.
- Informally, a restriction of a function f is the result of trimming its domain.
- $f(x) = x^2$,without any domain restriction, does not have an inverse function as it fails the horizontal line test.

$f^{-1}(x)$ is defined as the inverse function of f(x) if it consistently reverses the f(x) process. That is, if f(x) turns a into b, then $f^{-1}(x)$ must turn b into a. More concisely and formally, $f^{-1}(x)$ is the inverse function of f(x) if $f(f^{-1}(x)) = x$.

Informally, a restriction of a function f is the result of trimming its **domain**. More precisely, if S is any subset of X, the restriction of f to S is the function $f|S$ from S to Y such that $f|S(s) = f(s)$ for all s in S. If g is a restriction of f, then it is said that f is an extension of g. The overriding of $f: X \to Y$ by $g: W \to Y$ (also called overriding union) is an extension of g denoted as $(f \oplus g): (X \cup W) \to Y$. Its

Figure 5.6 Graph of y=x^2



This function fails the horizontal line test and therefore does not have an inverse.

graph is the set-theoretical union of the graphs of g and $f|X \setminus W$. Thus, it relates any element of the domain of g to its image under g, and any other element of the domain of f to its image under f. Overriding is an associative operation; it has the empty function as an identity element. If $f|X \cap W$ and $g|X \cap W$ are pointwise equal (e.g., the domains of f and g are disjoint), then the union of f and g is defined and is equal to their overriding union. This definition agrees with the definition of union for binary relations.

Domain restriction is important for inverse functions of exponents and logarithms because sometimes we need to find an unique inverse. For example, $f(x) = x^2$, without any domain restriction,

does not have an inverse function as it fails the horizontal line test *Figure 5.6*. But if we restrict the domain to be x>0 then we find that it passes the horizontal line test and therefore has an inverse function.

Section 2

Graphing Exponential Functions

https://www.boundless.com/algebra/exponents-and-logarithms/graphing-exponential-functions/

Basics of Graphing Exponential Functions

The exponential function $y = a \cdot b^x$] is a function that will remain proportional to its original value when it grows or decays.

KEY POINTS

- If the base, b, is greater than 1, then the function increases exponentially at a growth rate of b. This is known as exponential growth.
- If the base, b, is less than 1 (but greater than 0) the function decreases exponentially at a rate of b. This is known as exponential decay.
- If the base, b, is equal to 1, then the function trivially becomes $y = a$.

An exponential function is:

- A function that has exponents
- Is proportional to its original value whenever a quantity grows. This means that no matter what value is plugged in for the variable, it stays proportional to its original value.
- Is proportional to its original value whenever it decays. This means that if you were to take its derivative, its value would still be proportional to the original.

This can be shown with the function e^x. The derivative of e^x returns the same function, therefore it is always proportional to the original value.

Another example is the function $y = 2^x$. A graph of an exponential function becomes a curved line that steadily gets steeper, like in *Figure 5.7*. The basic formula for an exponential function is $y = a \cdot b^x$ where the constant a is the initial value or y-intercept of the function (0, a).

A good way to start a plot of the function is to choose numbers to plug in. For example, the function $y = 2^x$, choosing the x values 0, 1,

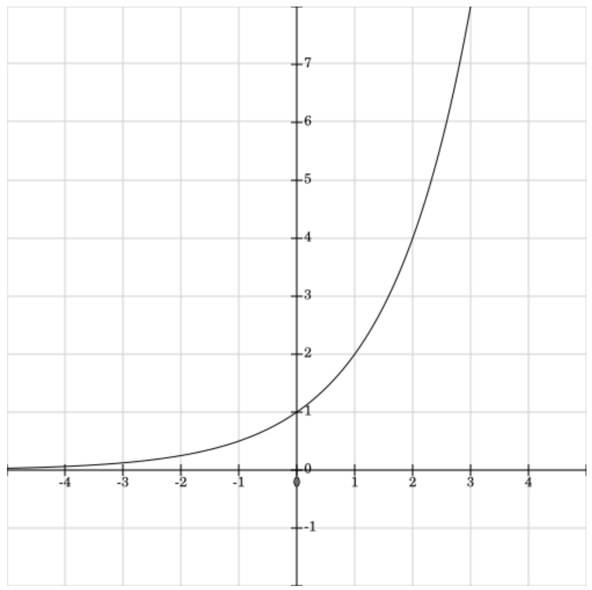


Figure 5.7 An Exponential Function

This is the graph of the exponential function 2^x.

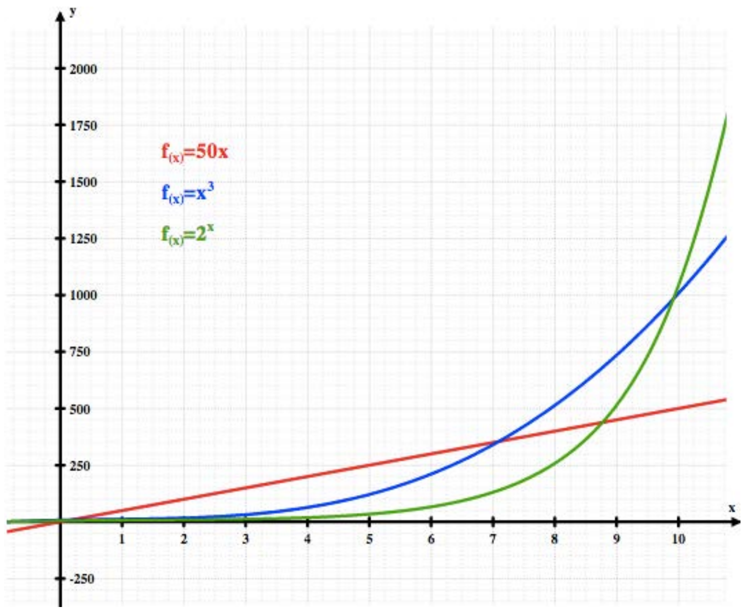


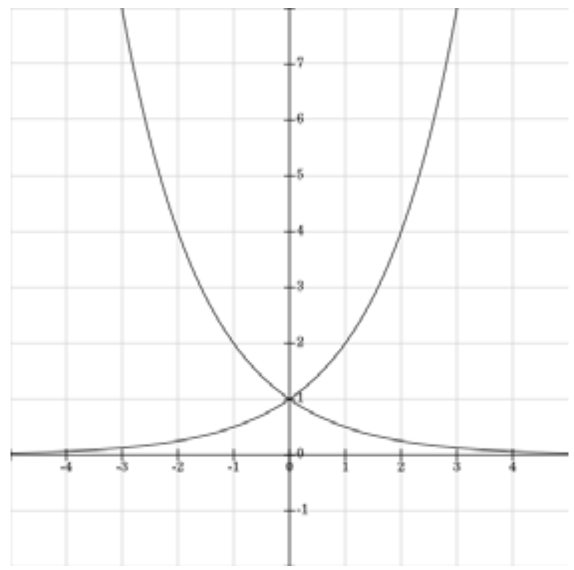
Figure 5.8 Exponential Functions vs. Others

This is a graph that shows how the exponential function will eventually grow much faster than any polynomial function.

2, 3, 4, and 5 you get the set of ordered pairs (0, 1), (1, 2), (2, 4), (3, 8), (4, 16), (5, 32). As you connect the points you will notice a smooth curve, as in *Figure 5.7*. You can see in *Figure 5.8* that exponential functions will grow very rapidly compared to polynomials.

If b is greater than 1, then it is known as exponential growth. If b is less than 1, but greater than 0, then it is known as exponential decay. The difference between the two can be seen by *Figure 5.9*, where the function $y = 2^x$ is growing exponentially, and increasing rapidly as x increases, whereas $y = (\frac{1}{2})^x$ is decaying exponentially, and decreases as x increases. In this case, the two are mirrors of each other across the y axis. Note that while neither function will ever reach 0, it will get infinitely close to 0 on one side. The line that

Figure 5.9 Exponential Growth and Decay



Graphed are two functions, one with exponential growth, the other exponential decay. The exponential growth increases rapidly as x increases, and exponential decay decreases rapidly as x increases.

the function approaches is referred to as the asymptote. The asymptote can be horizontal, vertical or oblique, depending on the function.

If $b = 1$, then the function becomes $y = a$, which is a horizontal line.

Problem Solving

Graphically solving problems with exponential functions allows visualization of sometimes complicated interrelationships.

KEY POINTS

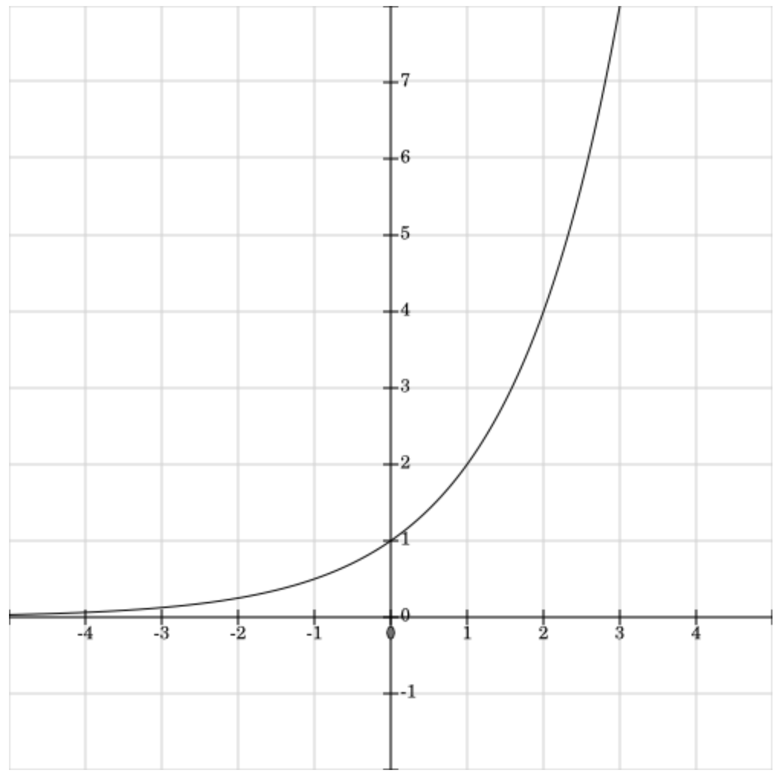
- Exponential functions are used in numerous applications, such as Biology, Economics, Finance, and many more.
- To model a general problem that you know is an exponential function, you only need two points to solve for the constants a and b.
- In the exponential function, $y = a \cdot b^x$, the constant a represents the initial value of the function, as at x = 0 the y value is a. The constant b is the rate of growth (or decay if less than 1) of the function.

The exponential function has numerous applications. Biology, Economics, and Finance are only just a few of the applications that use exponential functions.

Applications in Biology

In Biology, the number of microorganisms in a culture will increase exponentially until an essential nutrient is exhausted. Typically the first organism splits into two sister organisms, who then each split

Figure 5.10 An Exponential Function



This is a graph of an exponential function that is a representation of microorganism growth. In this case, the Y axis can be taken as the bacterial count, and the X axis is time. The resulting curve is nonlinear with respect to time -- it is exponential.

to form four, who split to form eight, and so on, which is the simple function $f(t) = 2^{t/x}$ where x is the time it takes for the organism to split (*Figure 5.10*). With this function, you can figure out how many organisms you will have after a certain amount of time. You can also determine how long it would take to reach a certain number of organisms with the help of the logarithm function, log(x).

A virus (for example SARS, or smallpox) typically will spread exponentially at first, if no artificial immunization is available. Each infected person can infect multiple new people.

Human population, if the number of births and deaths per person per year were to remain at current levels, is another example. According to the United States Census Bureau, over the last 100 years (1910 to 2010), the population of the United States of America is exponentially increasing at an average rate of one and a half percent a year (1.5%). This means that the **doubling time** of the American population (depending on the yearly growth in population) is approximately 50 years. A related constant is **half-life**, which is the amount of time required for a process to consume half the original quantity, or the time it takes a decaying substance to drop to half its original value.

Applications in Economics

Economic growth is expressed in percentage terms, implying **exponential growth**. For example, U.S. GDP per capita has grown at an exponential rate of approximately two percent per year for two centuries.

Applications in Finance

Compound interest at a constant interest rate provides exponential growth of the capital. Pyramid schemes or Ponzi schemes also show this type of growth resulting in high profits for a few initial investors and losses among great numbers of investors.

Solving General Problems

In general, a real world application that can be modeled with an exponential function is one that will continuously increase at an increasing rate based on the current value. For instance, in the microorganisms example, if the current value is 8, then the next time step will give 16, whereas if the current level is 1024, the next time step gives you 2048.

To figure out the values used to model a certain function, you need either an initial point (at x = 0), and any other point, or two general points if you do not have an initial point. For instance, let us say you have the two points (0, 4) and (3, 108). With the initial point (0, 4), you can plug it into the equation to get that a = 4. Now we know that the function takes the form $y = 4b^x$, and we need to next find b. With the second point we find that $108 = 4b^3$, and $b^3 = 27$, which means that $b = 3$. We now have the full equation $y = 4 \cdot 3^x$.

Now let us say that we have two general points, say (2, 12) and (5, 96). Plugging the first point into the equation gives us $12 = a \cdot b^2$, or $a = \frac{12}{b^2}$. Now with the second equation we have $96 = a \cdot b^5$, plugging in the value of a gives us $96 = \frac{12b^5}{b^2}$, or $b^3 = 8$, and so $b = 2$. Plugging

that back into the equation for a gives us $a = 3$. We now have our final equation of $y = 3 \cdot 2^x$.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/graphing-exponential-functions/problem-solving/

e

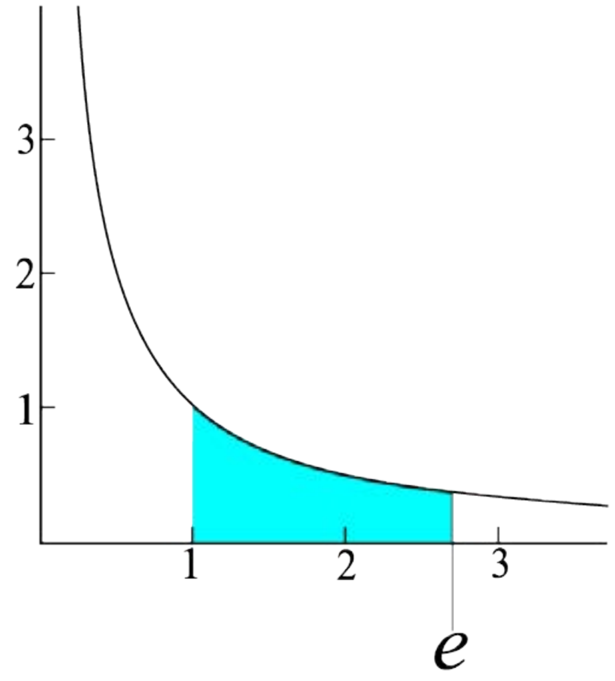
The number e is an important mathematical constant, approximately equal to 2.71828, that is the base of the natural logarithm, ln(x).

KEY POINTS

- The number e is defined in many ways, one of which is $\lim_{n \to \infty} (1 + \frac{1}{n})^n$, which was discovered through the idea of compound interest.
- The natural logarithm, written f(x) = ln(x), is the power to which e must be raised to obtain x.
- The constant can be defined in many ways; for example, e is the unique real number such that the derivative (slope of the tangent line) of the function e^x.
- More generally, an account that starts at $1 and offers an annual interest rate of R will, after t years, yield e^{Rt} dollars with continuous compounding.

The number **e** is an important mathematical constant, approximately equal to 2.71828, that is the base of the natural **logarithm**. It is $\lim_{n \to \infty} (1 + \frac{1}{n})^n$, an expression that arises in the study

Figure 5.11 The Constant e.



The area between the x-axis and the graph y = 1/x, between x = 1 and x = e is 1.

of compound interest, and can also be calculated as the sum of the infinite series: $e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots.$

The logarithm is the power to which a given base number must be raised in order to obtain x. For the natural logarithm, written ln(x), this is the power to which e must be raised to obtain x. For example, ln(e) = 1 and ln(1) = 0.

The constant can be defined in many ways; for example, e is the unique real number such that the derivative (slope of the tangent line) of the function $f(x) = e^x$ is e^x.

Sometimes called Euler's number after the Swiss mathematician Leonhard Euler, e is not to be confused with γ—the Euler–Mascheroni constant, sometimes called simply Euler's constant. The number e is also known as Napier's constant, but Euler's choice of this symbol is said to have been retained in his honor. The number e is of eminent importance in mathematics, alongside 0, 1, π and i. All five of these numbers play important and recurring roles across mathematics, and are the five constants appearing in one formulation of Euler's identity. Like the constant π, e is irrational: it is not a ratio of integers; and it is transcendental: it is not a root of any non-zero polynomial with rational coefficients.

Compound Interest

Jacob Bernoulli discovered this constant by studying a question about compound interest: An account starts with $1.00 and pays 100 percent interest per year. If the interest is credited once, at the end of the year, the value of the account at year-end will be $2.00. What happens if the interest is computed and credited more frequently during the year?

If the interest is credited twice in the year, the interest rate for each 6 months will be 50%, so the initial $1 is multiplied by 1.5 twice, yielding $1.00×1.52 = $2.25 at the end of the year. Compounding quarterly yields $1.00×1.254 = $2.4414..., and compounding monthly yields $1.00×(1+1/12)12 = $2.613035... If there are n compounding intervals, the interest for each interval will be 100%/n and the value at the end of the year will be $1.00×(1 + 1/n)n.

Bernoulli noticed that this sequence approaches a limit (the force of interest) with larger n and, thus, smaller compounding intervals. Compounding weekly (n = 52) yields $2.692597..., while compounding daily (n = 365) yields $2.714567..., just two cents more. The limit as n grows large is the number that came to be known as e; with continuous compounding, the account value will reach $2.7182818.... More generally, an account that starts at $1 and offers an annual interest rate of R will, after t years, yield e^{Rt} dollars with continuous compounding. (Here R is a fraction, so for 5% interest, R = 5/100 = 0.05).

Source: https://www.boundless.com/algebra/exponents-and-logarithms/graphing-exponential-functions/e/

Graphs of Exponential Functions, Base e

The function $f(x) = e^x$ is a basic exponential function with some very interesting properties.

KEY POINTS

- The function $f(x) = e^x$ is the only function whose derivative is itself. In other words, the slope of a tangent line to the graph at any point is equal to the y-value at that point.
- The exponential function arises whenever a quantity grows or decays at a rate proportional to its current value.
- The exponential function can be characterized as the power series: $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots$
- $\frac{d}{dx} e^x = \quad e^x$

The basic **exponential function**, sometimes referred to as the exponential function, is $f(x) = e^x$, where e is the number (approximately 2.718281828) such that the function e^x is its own derivative. The exponential function is used to model a relationship in which a constant change in the independent variable gives the same proportional change (i.e., percentage increase or decrease) in

the dependent variable. The function is often written as exp(x), especially when it is impractical to write the independent variable as a superscript. The exponential function is widely used in physics, chemistry, and mathematics.

The graph of $y = e^x$ is upward-sloping, and increases faster as x increases. The graph always lies above the x-axis, but can get arbitrarily close to it for negative x; thus, the x-axis is a horizontal **asymptote**. The slope of the **tangent** to the graph at each point is equal to its y-coordinate at that point.

The exponential function arises whenever a quantity grows or decays at a rate proportional to its current value. One such situation

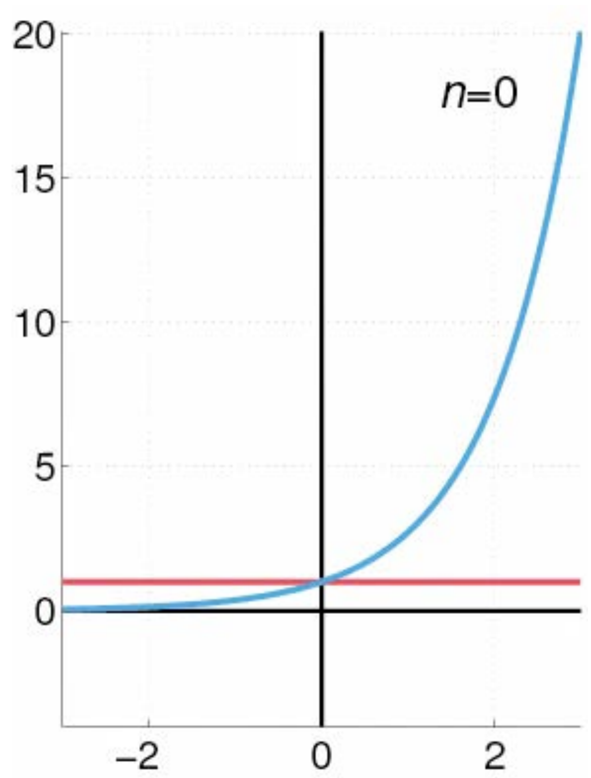


Figure 5.12 The Exponential Function as an Infinite Series

This is an animation showing the exponential function (in blue) and the first n+1 terms of the power series (in red) that make up the exponential function.

is continuously compounded interest, which is exactly what led Jacob Bernoulli in 1683 to the number now known as e. Later, in 1697, Johann Bernoulli studied the calculus of the exponential function.

Formal Definition

The exponential function e^x can be characterized in a variety of equivalent ways. In particular it may be defined by the following power series:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$$

This can be seen in *Figure 5.12*. As you add more and more of the power series, you get closer and closer to the exponential function.

The function may also be defined as the following limit:

$$e^x = \lim_{n\to\infty} (1 + \frac{x}{n})^n$$

The Graph

The graph of $f(x) = e^x$, as shown in *Figure 5.13*, is like any other exponential function, where it has important values (0, 1) and (1, e). There are some very interesting properties about the exponential function's graph that no other function has. The main property of

note is that the slope of the tangent line to the graph at any point is equal to the y-value at that point, as seen in [Figure 5.14](#). Or, as stated before, the derivative of e^x is e^x.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/graphing-exponential-functions/graphs-of-exponential-functions-base-e/

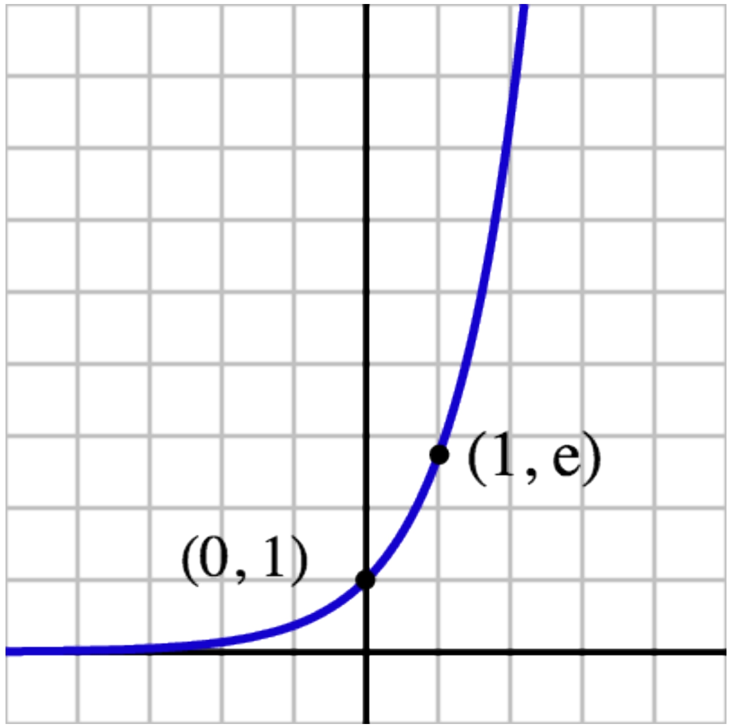


Figure 5.13 The Exponential Function

The graph of the natural exponential function with base e, y=ex.

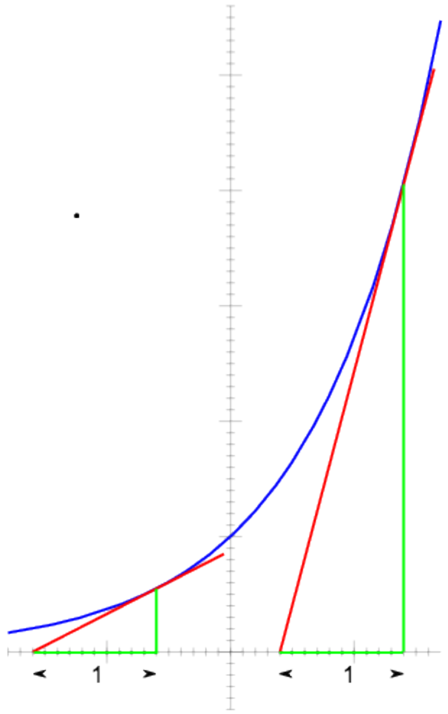


Figure 5.14 Tangents to the Exponential Function

The tangent to the exponential function with base e at any point is equal to the y-value at that point.

Section 3

Graphing Logarithmic Functions

https://www.boundless.com/algebra/exponents-and-logarithms/graphing-logarithmic-functions/

Logarithmic Functions

The logarithm of a number is the exponent by which another fixed value, the base, has to be raised to produce that number.

KEY POINTS

- The inverse of the logarithmic operation is exponentiation.
- The logarithm is commonly used in many fields: that with base 2 in computer science, that with base e in pure mathematics, and that with base 10 in natural science and engineering.
- The logarithm of a product is the sum of the logarithms of the factors.

In its simplest form, a logarithm is an **exponent**. Taking the **logarithm** of a number, one finds the exponent to which a certain value, known as a base, is raised to produce that number once more.

Logarithms have the following structure:

$$log_b(x) = c$$

where b is known as the base, c is the exponent to which the base is raised to afford x.

Consider the following logarithm, for example:

$$log_3(243) = 5$$

The left side of the equation states that the right will be the exponent to which 3 is raised to yield 243 and indeed, $3^5=243$.

The explanation of the previous example reveals the inverse of the logarithmic operation: **exponentiation**. Starting with 243, if we take its logarithm with base 3, then raise 3 to the logarithm, we will once again arrive at 243.

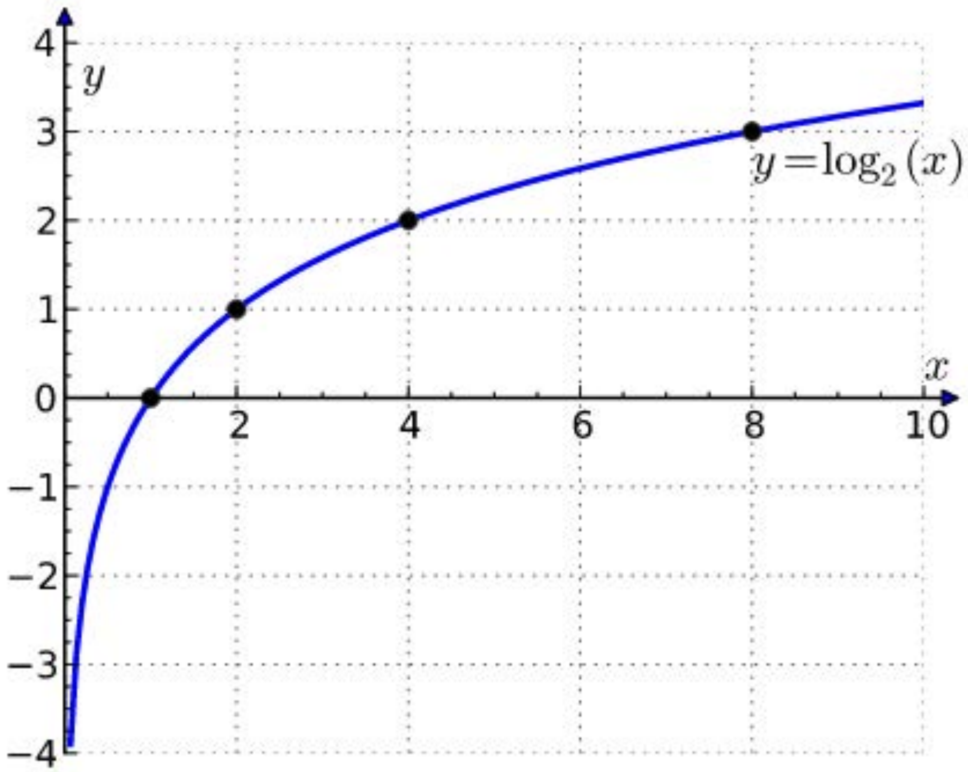
Logarithms are most commonly used to simplify complex calculations that involve high-level exponents. In science and engineering, the logarithm with base 10 is commonly used.

In pure mathematics, the logarithm with base e (≈ 2.718) is often applicable. In computer science, the binary logarithm, with base 2, is commonly used.

In natural science and engineering, the logarithm with base 10 is frequently used. In chemistry, for example, pH and pKa are used to simplify concentrations and dissociation constants, respectively, of high exponential value.

The purpose is to bring wide-ranging values into a more manageable scope. A dissociation constant may be smaller than

Figure 5.15 Graph of Binary Logarithm



The function's slope decreases with increasing x, thus containing its vertical growth. More importantly, the scale is greatly reduced, allowing for easy visualization of large values of y. In many cases, by taking the logarithms of exponential equations, the equations become linear, and plotting the appropriate logarithms can help solve for roots, slopes, and intercepts.

10^{10}, or higher than 10^{-50}. Taking the logarithm of each brings the values into a more comprehensible scope (10 to -50) (*Figure 5.15*).

Special Logarithms

Any number can be used as the base of a logarithm but certain bases (10, e, and 2) have more widespread applications than others.

KEY POINTS

- Logarithms with base equal to 10 are called common logarithms. They are most applicable in physical and natural sciences and engineering.
- Logarithms with base equal to e are called natural logarithms. They are most applicable in pure mathematics.
- Logarithms with base equal to 2 are called binary logarithms. They are most applicable in computer science.

Not all **logarithms** are equally useful in practice. Some, depending on their base, are more useful than others in a given field of study.

Out of the infinite number of possible bases, three stand out as particularly useful.

Logarithms with base 10 are called common logarithms. They are named so because they are widely used, and in fact the common logarithm of x can be denoted as:

$$log(x)$$

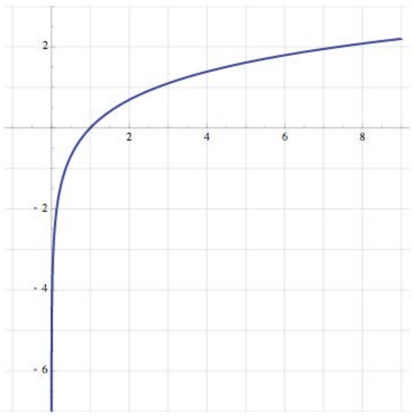


Figure 5.16 Graph of the Natural Logarithm

The natural logarithm is the logarithm with base equal to Euler's number, e.

Note that the base is not specified; it is implied.

Common logarithms are often used in physical and natural sciences and engineering. For example, the magnitude of an earthquake (M) can be determined based on the logarithm of an intensity measurement from a seismograph (I):

$$M = log(\frac{I}{I_0})$$

where Io is a constant.

Logarithms with base e are called natural logarithms. They are so commonly used that they are have a unique symbol. The natural logarithm with base x is denoted as:

$$ln(x)$$

Like common logarithms, natural logarithms are not expressed with a base. The base of e is implicit in their symbol (*Figure 5.16*)

Natural logarithms are often found in physical sciences and pure math. For example, the entropy (S) of a system can be calculated from the natural logarithm of the number of possible microstates (W) the system can adopt:

$$S = k * ln(W)$$

where k is a constant.

Logarithms with base 2 are called binary logarithms. The binary logarithm of x is commonly written as:

$$ld(x)$$

or

$$lb(x)$$

Binary logarithms are useful in any application that involves the doubling of a quantity, and particularly in computer science with the use of integral parts.

Converting Between Exponential and Logarithmic Equations

Logarithmic and exponential forms are closely related, and an equation in either form can be freely converted into the other.

KEY POINTS

- The logarithmic and exponential operations are inverses.
- If given an exponential equation, one can take the natural logarithm to isolate the variables of interest, and vice versa.
- Converting from logarithmic to exponential form can make for easier equation solving.

Because **logarithmic** and exponential functions are inverses of one another, one can be converted into the form of the other.

Logarithmic equations have the form:

$$y = log_b(x)$$

where b is a base (of defined value) and x and y are the independent and dependent variables, respectively.

Exponential equations have the form:

$$x = b^y$$

where x and y are the dependent and **independent variables**, respectively, and b is of defined value (Figure 2).

If b is the same in both the above equations, the x and y from the logarithmic equation are respectively equal to the x and y in the exponential equation.

Thus, if given the exponential equation:

$$4^3 = 64$$

One can convert it to logarithmic form:

$$lob_4(64) = 3$$

And if given the logarithmic equation:

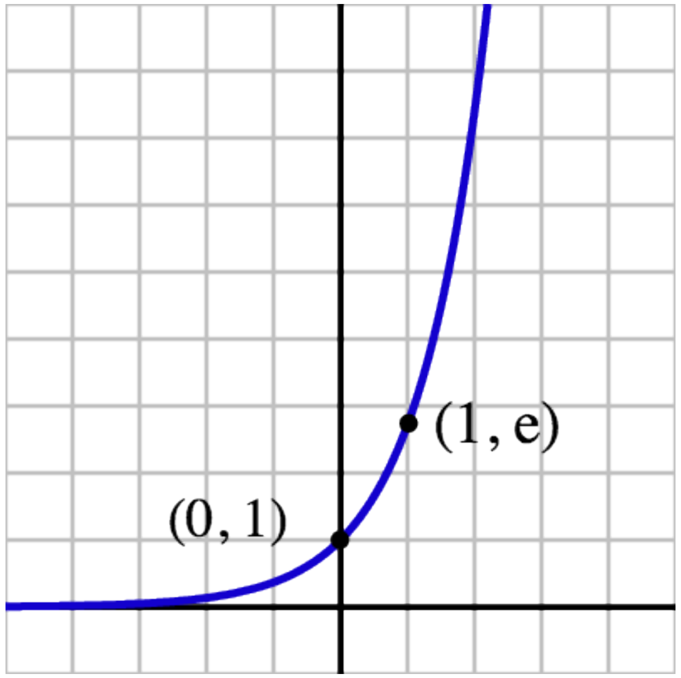
$$log_7(2401) = 4$$

One can convert to exponential form:

$$7^4 = 2401$$

Conversion from logarithmic to exponential form can help one solve otherwise difficult equations.

Figure 5.17 The Exponential Function



The natural exponential function ex.

Consider the following equation:

$log_6(x - 2) = 3$

Exponentiating each side gives:

$6^{log_6(x-2)} = 6^3$

As inverses, the exponential and logarithm cancel, giving:

$x - 2 = 216$

Thus:

$x = 218$

Source: https://www.boundless.com/algebra/exponents-and-logarithms/graphing-logarithmic-functions/converting-between-exponential-and-logarithmic-equations/

Natural Logarithms

The natural logarithm is the logarithm to the base e, where e is an irrational and transcendental constant approximately equal to 2.718281828.

KEY POINTS

- The natural logarithm is the logarithm with base equal to e.
- Also known as Euler's number, e is an irrational number that often appears in natural relationships in pure math and science.
- The number e and the natural logarithm have many applications in calculus, number theory, differential equations, complex numbers, compound interest, and more.

The logarithm of a number is the exponent by which another fixed value, the base, has to be raised to produce that number. The **natural logarithm** is the logarithm with base equal to **e** (*Figure 5.18*).

Also known as Euler's number, e is an irrational number representing the limit of:

$(1 + 1/n)^n$

as n approaches infinity. In other words, e is the sum of 1 plus 1/1 plus 1/(1*2) plus 1/(1*2*3), and so on.

The number e has many applications in calculus, number theory, differential equations, complex numbers, compound interest, and more. It also is extremely useful as a base in logarithms; so useful that the logarithm with base e has its own name (natural logarithm) and symbol. Here is the proper notation for the natural logarithm of x:

$ln(x)$

The natural logarithm is so named because unlike 10, which is given value by culture and has minimal intrinsic use, e is an extremely interesting number that often "naturally" appears, especially in calculus.

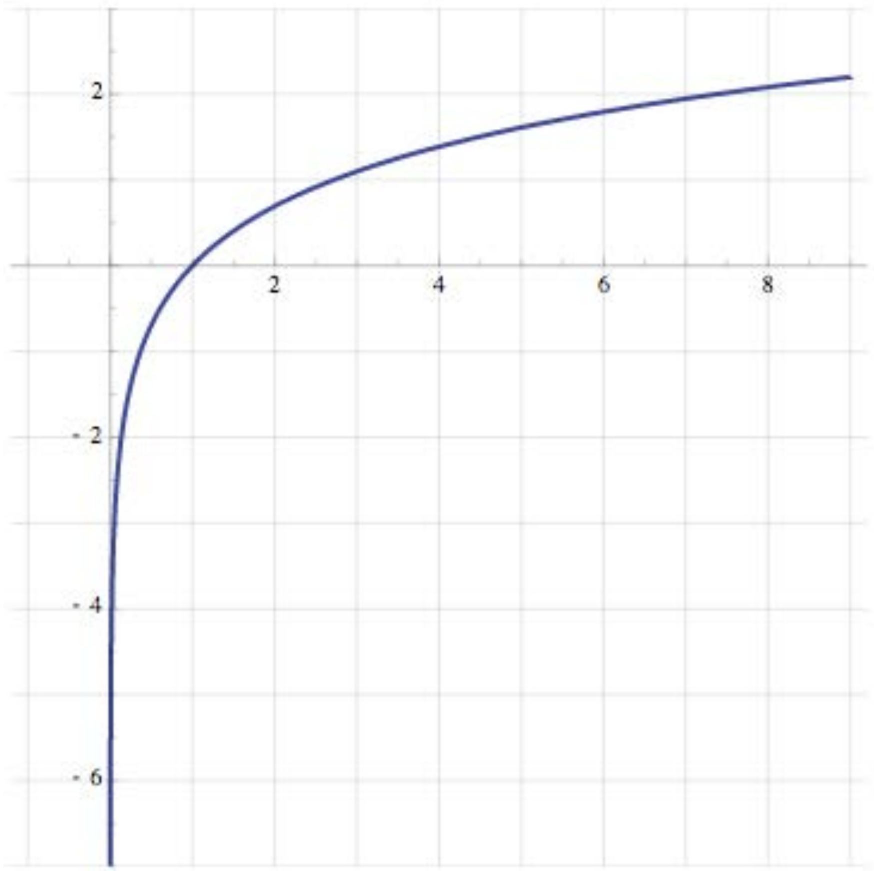
The inverse of the natural log appears, for example, upon differentiating a logarithm of any base:

$$\frac{d}{dx} log_b(x) = \frac{1}{x * ln(b)}$$

Outside of calculus, the natural logarithm can be used to relate 1, e, i, and π, four of the most important numbers in mathematics:

$ln(-1) = i\pi$

Figure 5.18 Graph of the Natural Logarithm



The function slowly grows to positive infinity as x increases and rapidly goes to negative infinity as x approaches 0 ("slowly" and "rapidly" as compared to any power law of x); the y-axis is an asymptote.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/graphing-logarithmic-functions/natural-logarithms/

Changing Logarithmic Bases

The base of a logarithm can be converted to another value through a simple, one-step process.

KEY POINTS

- If a logarithm and/or its base are not whole numbers, evaluation can be near-impossible.
- One can change the base of a logarithm by expressing it as the quotient of two logarithms with a common, same base.
- Changing a logarithm's base to 10 makes it much simpler to evaluate; it can be done on a calculator.

So long as a **logarithm**, its **base**, and the number upon which it operates are all whole numbers, one can evaluate logarithms manually with minimal difficulty.

When decimals are involved, however, it can become exceedingly difficult to evaluate a logarithm. Let's consider:

$log_4(9)$

We can easily determine that the above will simplify to a number between one and two, because $4^1 = 4$ and $4^2 = 16$. The exact value, however, is not so easily determined.

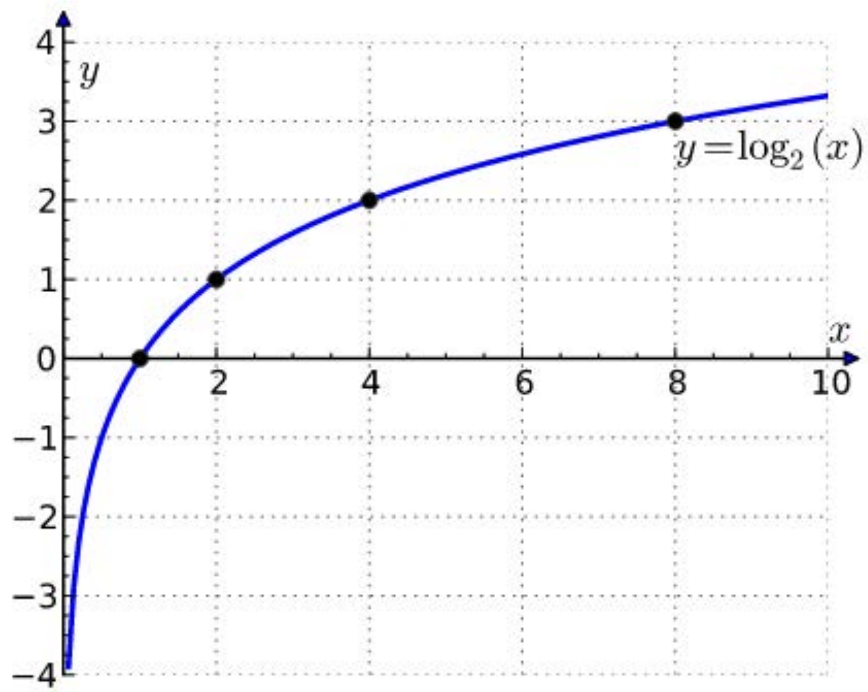
Not all calculators have logarithm functions and, those that do almost always have a base of 10. Fortunately, any logarithm can be converted into a logarithm of equal value with a different base. The formula for this transformation is:

$$log_a(x) = \frac{log_b(x)}{log_b(a)}$$

where a is the original base and b is the desired base.

Revisiting the example above, we can change the base from 4 to 10, which can be input into a calculator.

Figure 5.19 Graph of Binary Logarithm



This function can be expressed as the quotient of log (x) and log (2), both with a base of 10.

$$log_4(9) = \frac{log_{10}(9)}{log_{10}(4)}$$

The left side of the equation is extremely difficult to calculate manually, and would be impossible to find on most calculators. However, the quotient of logarithms with base equal to 10 can easily be found on a scientific calculator (*Figure 5.19*).

Graphs of Logarithmic Functions

Logarithms can be graphed manually or electronically with points generally determined via a calculator or table.

KEY POINTS

- The logarithmic graph is similar in shape to the square root graph, but with a vertical asymptote as x approaches 0 from the right.
- With use of defined numbers, logarithmic graphs can be shifted horizontally and/or vertically.
- When graphing logarithms, it is usually advisable to use several x values between the vertical asymptote and the next integer. After the first integer, it is usually best to make points less and less frequently.

At first glance, the graph of the logarithmic function can easily be mistaken for that of the square root function.

Both the square root and logarithmic functions have a domain limited to x values greater than 0. However, the **logarithmic function** has a vertical asymptote descending towards negative ∞ as x approaches 0, whereas the square root reaches a minimum y value of 0.

The logarithmic graph begins with a steep climb after x=0, but stretches more and more horizontally, its slope ever-decreasing with increasing x. Thus, the graph of the logarithmic function shows, as expected, an inverse relationship to the graph of the exponential.

Graphing logarithmic functions can be performed manually or with a calculator. In either case, except if graphing a natural logarithm on a calculator, one must first convert the logarithm's base to 10. This is easily done using a table or an electronic calculator, converting the logarithm to its original form, then taking the log in base 10. Base 10 is simple to understand in that the numbering system is based on 10, so working with powers of 10 is "almost" second nature.

The logarithm to base b = 10 is called the common logarithm and has many applications in science and engineering. The natural logarithm has the constant e (≈ 2.718) as its base. Its use is widespread in pure mathematics, especially calculus. The binary logarithm uses base b = 2 and is prominent in computer science.

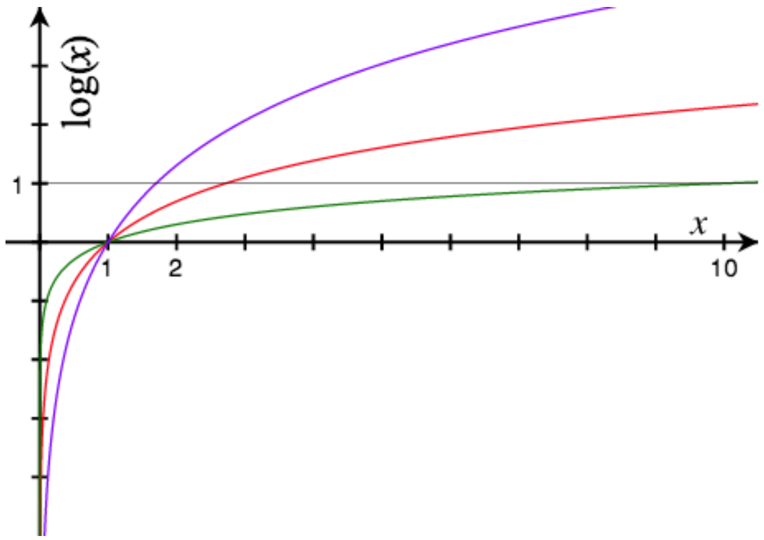
Logarithms were introduced by John Napier in the early 17th century as a means to simplify calculations. They were rapidly

adopted by navigators, scientists, engineers, and others to perform computations more easily, using slide rules and logarithm tables. Tedious multi-digit multiplication steps can be replaced by table look-ups and simpler addition because of the fact — important in its own right — that the logarithm of a product is the sum of the logarithms of the factors:

The present-day notion of logarithms comes from Leonhard Euler, who connected them to the exponential function in the 18th century.

Logarithmic scales reduce wide-ranging quantities to smaller scopes. For example, the decibel is a logarithmic unit quantifying

Figure 5.20 Logarithms



All logarithmic functions will have the same general shape, but their graphs will vary depending on base and any coefficients inserted into the equation. In this figure, red is ln(x) (which is log base e), green is log(x) base 10, purple is log(x) base 1.7.

sound pressure and voltage ratios. In chemistry, pH is a logarithmic measure for the acidity of an aqueous solution. Logarithms are commonplace in scientific formulae, and in measurements of the complexity of algorithms and of geometric objects called fractals. They describe musical intervals, appear in formulae counting prime numbers, inform some models in psychophysics, and can aid in forensic accounting.

Once the logarithm is converted to a base of 10, one can create a table of points using x values and calculating their corresponding y values.

Generally it is best to use several x values between 0 and 1 (0.125, 0.25, 0.5, and 1) to capture the asymptotic rise, and then double x values (2, 4, 8, 16) for every point thereafter.

Graphs of logarithmic functions can be shifted horizontally or vertically by incorporating other values inside and outside the log brackets (*Figure 5.20*):

$y = log(x + a)$

where a is a certain defined number, shifts the graph horizontally. If a is a positive number, the graph shifts to the left; if a is negative, the graph shifts to the right. Such a shift would change the x values of interest in determining points to graph.

$$y = log(x) + a$$

where a is a certain defined number, shifts the graph vertically. If a is positive, the graph shifts upward; if a is negative, the graph shifts downward.

Solving Problems with Logarithmic Graphs

Some functions with rapidly changing shape are best plotted on a scale that increases exponentially; such scales make up logarithmic graphs.

KEY POINTS

- Logarithmic graphs use logarithmic scales, in which the values differ exponentially. For example, instead of including marks at 0, 1, 2, and 3, a logarithmic scale may include marks at 0.1, 1, 10, and 100, each an equal distance from the previous and next.
- Logarithmic graphs allow one to plot a very large range of data without losing the shape of the graph.
- Logarithmic graphs make it easier to interpolate in areas that may be difficult to read on linear axes. For example, if the plot of y=x5 is scaled to show a very wide range of y values, the curvature near the origin may be indistinguishable on linear axes. It is much clearer on logarithmic axes.

Many mathematical and physical relationships are functionally dependent on high-order variables.

Consider the Stefan-Boltzmann law, which relates the power (j*) emitted by a black body to temperature (T).

$J* = \sigma T^4$.

On a standard graph, this equation can be quite unwieldy. The fourth-degree dependence on temperature means that power increases extremely quickly. The fact that the rate is ever-increasing (and steeply so) means that changing scale is of little help in making the graph easier to interpret.

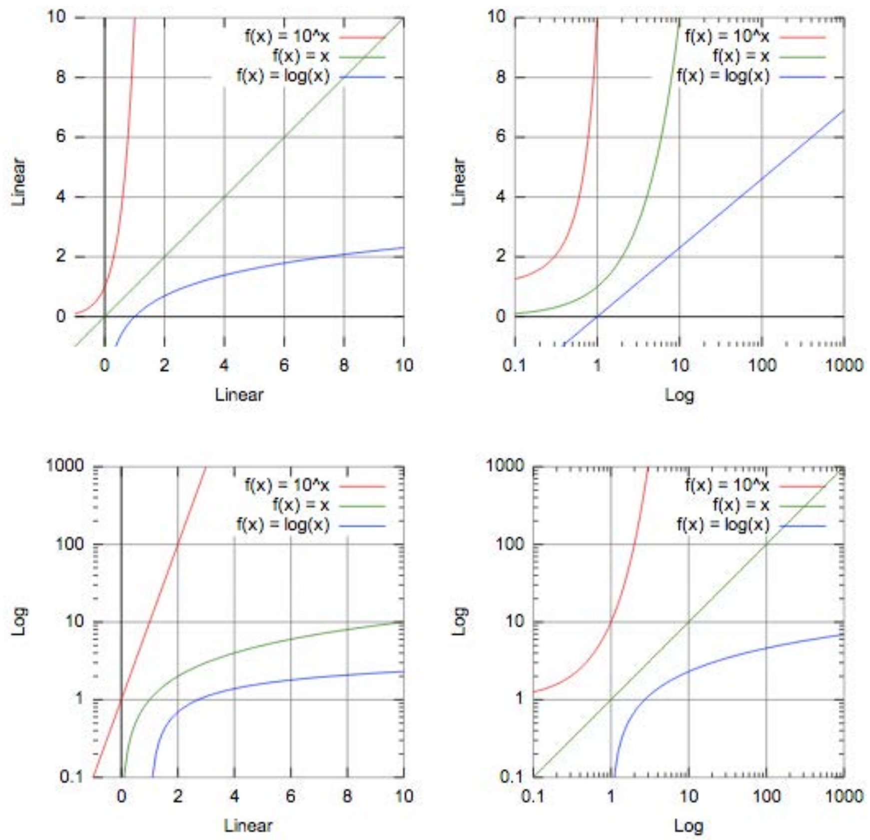
For very steep functions, it is possible to plot points more smoothly while retaining the integrity of the data: one can use a graph with a logarithmic scale (*Figure 5.21*).

The primary difference between the logarithmic and linear scales is that, while the difference in value between linear points of equal distance remains constant (that is, if the space from 0 to 1 on the scale is 1cm on the page, the distance from 1 to 2, 2 to 3, etc., will be the same), the difference in value between points on a logarithmic scale will change exponentially. A logarithmic scale will start at a certain power of 10, and with every unit will increase by a power of 10.

Thus, if one wanted to convert a linear scale (with values 0-5) to a logarithmic scale, one option would be to replace 0, 1, 2, 3, 4, and 5 with 0.001, 0.01, 0.1, 1, 10, and 100, respectively. Between each major value on the logarithmic scale, the hashmarks become increasingly closer together with increasing value. For example, in the space between 1 and 10, the 8 and 9 are much closer together than the 2 and 3.

The advantages of using a logarithmic scale are twofold. Firstly, doing so allows one to plot a very large range of data without losing the shape of the graph. Secondly, it allows one to **interpolate** at

Figure 5.21 Logarithmic Scale



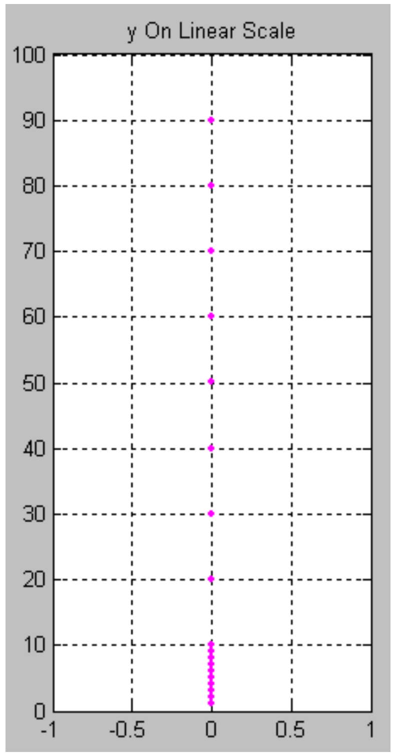
The graphs of functions f(x)=10^x, f(x)=x, and f(x)=log(x) on four different coordinate plots. Note how each function changes shape on each set of coordinates.

any point on the plot, regardless of the range of the graph (*Figure 5.23*). Similar data plotted on a linear scale is less clear (*Figure 5.22*).

A key point about using logarithmic graphs to solve problems is that they expand scales to the point at which large ranges of data make more sense. In logarithms, the product of numbers is the sum of their logarithms. In the equation mentioned above, plotting j vs. T would generate the expected curve, but the scale would be such that minute changes go unnoticed and the large scale effects of the relationship dominate the graph—it is so big that the "interesting areas" won't fit on the paper on a readable scale.

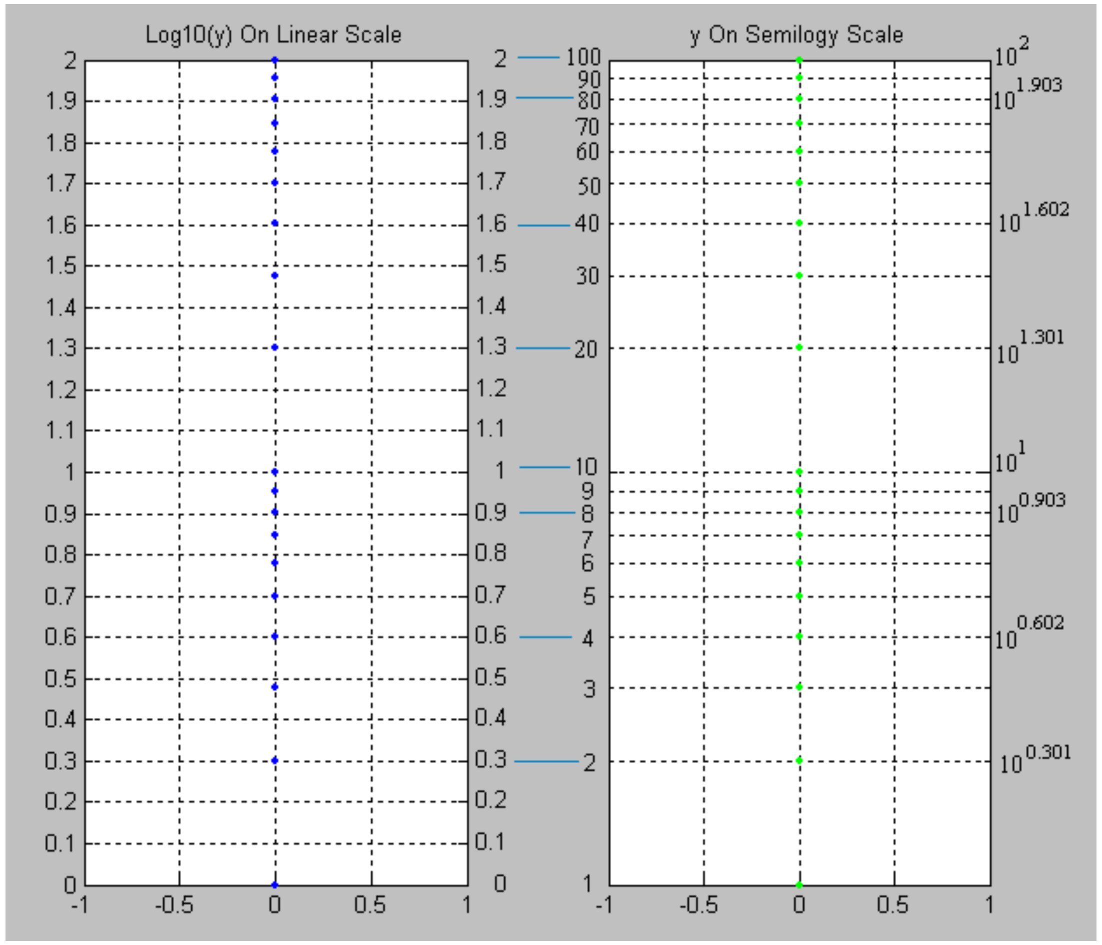
Taking logarithms, however, one ends with $log j = 4 * log(\sigma T) = 4 * log\sigma + 4 log T$. This is the equation of a straight line with a slope of $log T$ and a "y intercept" of $4 \log \sigma$. Plotting a straight line as indicated simplifies the interpretation.

Figure 5.22 Points of log(y) on a Linear Scale



Notice how values of y less than 10 are indistinguishable.

Figure 5.23 Graph of log(y) on a Semi-Log Scale



Both plots capture y well for their respective ranges, but note how easily distinguishable the points are in both the lower and higher areas.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/graphing-logarithmic-functions/solving-problems-with-logarithmic-graphs/

Section 4

Properties of Logarithmic Functions

https://www.boundless.com/algebra/exponents-and-logarithms/properties-of-logarithmic-functions/

Logarithms of Products

A useful property of logarithms states that the sum of two logarithms of factors is equal to the logarithm of the factors' product.

KEY POINTS

- The logarithm of a number is the exponent by which another fixed value, the base, has to be raised to produce that number.
- The logarithm of a product is the sum of the logarithms of the factors. In addition, the sum of the logarithms of two number is equal to the logarithm of the product of those two numbers.
- The product rule does not apply when the base of the two logarithms are different.

The logarithm of a number is the **exponent** by which another fixed value, the base, has to be raised to produce that number. For example, the logarithm of 1000 to base 10 is 3, because 1000 is 10 to the power 3: 1000 = 10 × 10 × 10 = 103. More generally, if x = by, then y is the logarithm of x to base b, and is written $y = log_b(x)$, so $log_{10}(1000) = 3$. The shape of the graphs that logarithms take can be viewed in *Figure 5.24*.

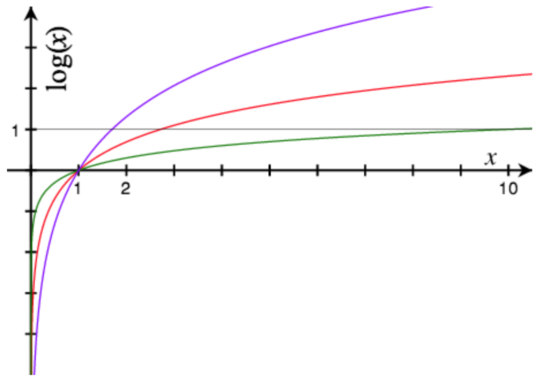


Figure 5.24
Logarithms
The graphs of logarithms of different bases have the same general shape but different curvatures.

Logarithms were introduced by John Napier in the early 17th century as a means to simplify calculations. Logarithms were rapidly adopted by navigators, scientists, engineers, and others to perform computations more easily, using slide rules and logarithm tables. Tedious multi-digit multiplication steps can be replaced by table look-ups and simpler addition, because of the fact that the logarithm of a product is the sum of the logarithms of the factors:

$$logb(xy) = logb(x) + logb(y)$$

Of course, the reverse is true also. The sum of two logarithms of factors is equal to the logarithm of the factors' product. This is a very useful property of logarithms, because it can sometimes simplify more complex equations. However, it is also important to note that this property is only true when the two logarithms share the same base. An example of this would be:

$$logb(x) + logc(y) = logb(xy)$$

This equation is not true, of course, in a situation such as b = c.

Logarithms of Powers

A simplifying principle of logarithms is that the logarithm of the p-th power of a number is p times the logarithm of the number.

KEY POINTS

- The logarithm of a product is the sum of the logarithms of the factors.
- An exponent, p, signifies that a number is being multiplied by itself p number of times. Because the logarithm of a product is the sum of the logarithms of the factors, the logarithm of a number, a, to an exponent, p, is the same as the logarithm of a added together p times.
- The logarithm of a added together p times is the same as $p \cdot log_b(a)$, where b is an arbitrary base.

The **logarithm** of a number is the **exponent** by which another fixed value, the **base**, has to be raised to produce that number. For example, the logarithm of 1000 to base 10 is 3,written: $log_10(1000) = 3$. *Here, I have chosen to write the 10, to show where you would denote the base, however if you are working with a base of 10, you do not have to denote that. If it is left blank, you can assume you are looking for the base10. This is true because:

$10^3 = 10 * 10 * 10 = 1000$. More generally, if x = by, then y is the logarithm of x to base b, and is written $log_b x$.

Logarithms are used commonly to measure earthquakes, distances of stars, economics, and throughout the scientific world.

Logarithms were introduced by John Napier in the early 17th century as a means to simplify calculations. They were adopted by navigators, scientists, engineers, and others to perform computations more easily, using slide rules and logarithm tables. Tedious multi-digit multiplication steps can be replaced by table look-ups and simpler addition because of the fact that the logarithm of a product is the sum of the logarithms of the factors:

$$log_b(ac) = log_b(a) + log_b(c)$$

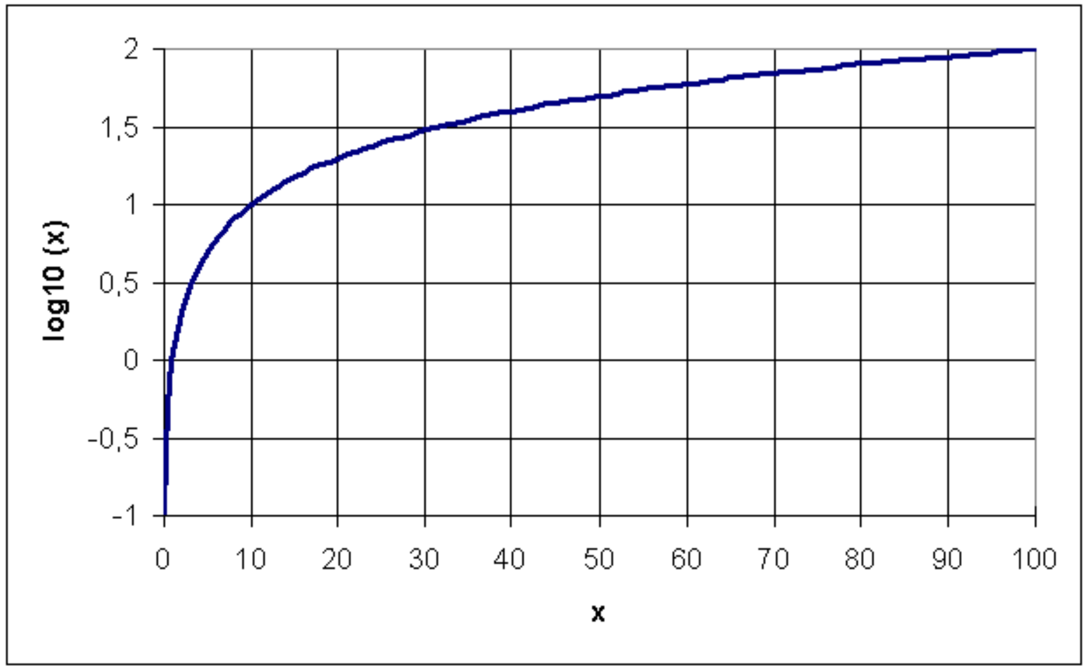
The logarithm of the p-th power of a number is p times the logarithm of the number itself.

$$log_b(a)^p = p * log_b(a)$$

This can be proved by using the fact that the log of a product is equal to the sum of the logs of the two factors. Because an exponent just indicates the the number is multiplied by itself many times, we can write:

$$log_b(a)^p = log_b(a * a * a * \cdots * a_p)$$

Figure 5.25 Graph of Common Logarithm



The graph of a common logarithm takes a characteristic shape.

Where a is multiplied by itself p times. This equation can be written by following the product rule:

$$log_b(a) + log_b(a) + log_b(a) + \cdots + log_b(a_p)$$

Where $log_b a$ is added together with itself p times. This, of course, by the simple rules of algebra, can be written as:

$$p * \log_b(a)$$

Thus, we have demonstrated the rule of the logarithm of powers.

Logarithms of Quotients

The logarithm of the ratio or quotient of two numbers is the difference of the logarithms and can be proven using the first law of exponents.

KEY POINTS

- The logarithm of a number is the exponent by which another fixed value, the base, has to be raised to produce that number.
- A basic idea in logarithmic math is that the logarithm of a product is the sum of the logarithms of the factors.
- A similar idea of the law of products is that the logarithm of the ratio or quotient of two numbers is the difference of the logarithms.

The logarithm of a number is the **exponent** by which another fixed value, the base, has to be raised to produce that number. For example, the logarithm of 1000 to base 10 is 3, because 1000 is 10 to the power 3: 1000 = 10 × 10 × 10 = 103. More generally, if x = by, then y is the logarithm of x to base b, and is written y = logb(x), so log10(1000) = 3 (*Figure 5.26*).

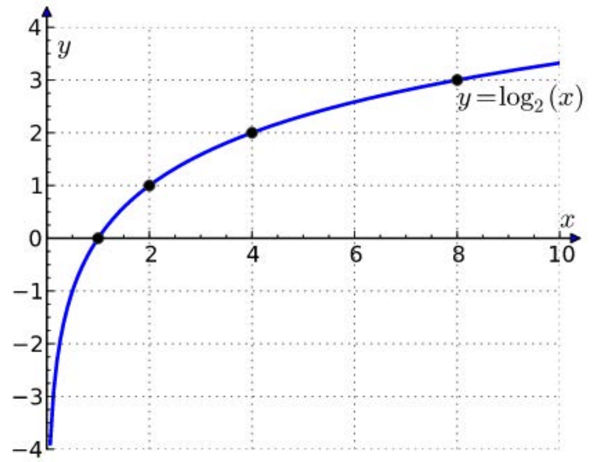
The logarithm of a product is the sum of the logarithms of the factors:

$log_b(xy) = log_b x + log_b y$

Similarly, the logarithm of the ratio of two numbers is the difference of the logarithms.

$log_x(a/b) = log_x a - log_x b$

Figure 5.26 Graph of Binary Logarithm



The graph of the logarithm to base 2 crosses the x axis (horizontal axis) at 1 and passes through the points with coordinates (2, 1), (4, 2), and (8, 3). For example, log2(8) = 3, because 23 = 8. The graph gets arbitrarily close to the y axis, but does not meet or intersect it.

To prove this, let $m = log_x a$.

Rewrite the above expression as an exponent. $x^m = a$ ($log_x a$ asks " x to what power is a ?" And the equation answers: "x to the m is a.")

Let $n = log_x b$. Thus, $x^n = b$.

If we replace a and b based on the previous equations, we get:

$log_x(a/b) = log_x(x^m/x^n)$

This can be further simplified to:

$log_x(x^m/x^n) = log_x(x^{m-n})$

Which, using the first law of exponents, can be written as:

$log_x(x^m/x^n) = m - n$

This is the key step. Therefore, it can be seen that the properties of logarithms come directly from the laws of exponents. Replacing m and n with what they were originally defined as results in this equation:

$log_x(a/b) = log_x a - log_x b$

Hence, the previous problem has been proven.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/properties-of-logarithmic-functions/logarithms-of-quotients/

Solving General Problems with Logarithms and Exponents

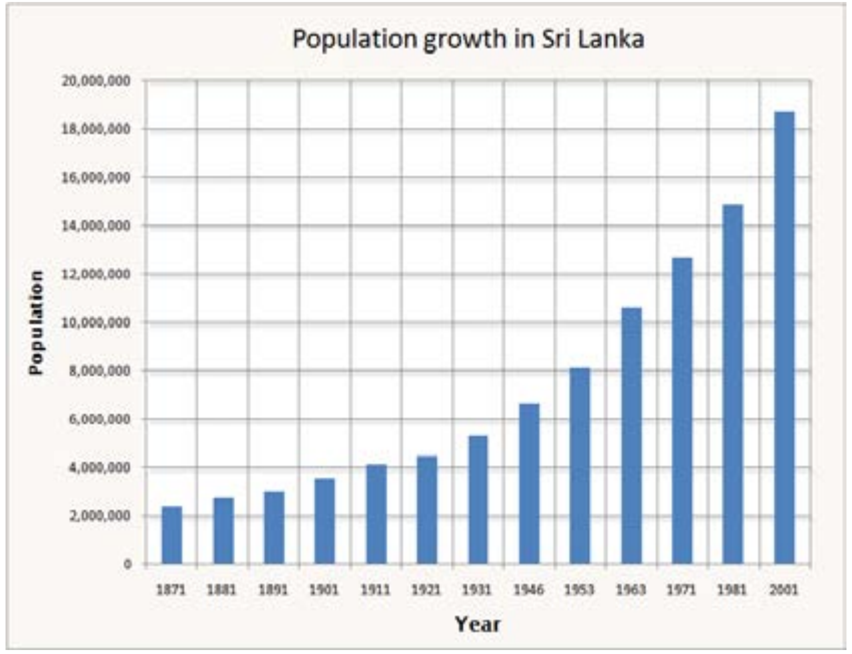
Logarithms are useful for solving equations that require an exponential term, like population growth.

KEY POINTS

- The logarithm of a number is the exponent by which another fixed value, the base, has to be raised to produce that number.
- Both root equations and logarithm equations can be rewritten as exponent equations.
- Logarithms are therefore useful in solving equations that require solving for an exponential term, such as those involving population growth.

The logarithm of a number is the exponent by which another fixed value, the base, has to be raised to produce that number. For example, the logarithm of 1000 to base 10 is 3, because 1000 is 10 to the power 3: 1000 = 10 × 10 × 10 = 103. More generally, if $x = b^y$ then y is the logarithm of x to base b, and is written $y = log_b x$, so log10(1000) = 3.

Figure 5.27 Population Growth in Sri Lanka



This population growth graph shows that it grows exponentially with time.

Logarithms have several applications in general math problems. To look at it most simply, both root equations and logarithm equations can be rewritten as exponent equations.

$\sqrt{9} = 3$ can be rewritten as $3^2 = 9$. These two equations are the same statement about numbers, written in two different ways. $\sqrt{9}$ asks the question "What number squared is 9?" So the equation $\sqrt{9} = 3$ asks this question, and then answers it: "3 squared is 9."

We can rewrite logarithm equations in a similar way. Consider this equation:

$$log3\frac{1}{3} = -1$$

If you are asked to rewrite that logarithm equation as an exponent equation, think about it this way. The left side asks: “3 to what power is (1/3)?” And the right side answers: “3 to the −1 power is (1/3).” $3^{-1} = 1/3$.

These two equations, $log_3 1/3 = -1$ and $3^{-1} = 1/3$, are two different ways of expressing the same numerical relationship.

EXAMPLE

A city grows 5% every 2 years. How long will it take for the city to triple its size?

Step 1. Use the growth formula:

$A = P(1 + i)^n$

Assume P = x. Since we want to know when the population will be tripled, A = 3x. For this example n represents a period of 2 years, therefore the n is halved for this purpose.

Step 2. Substitute information given into formula:

1. $3 = (1.05)^{n/2}$
2. $log 3 = n/2 * log 1.05$
3. $n = 2 log 3 / log 1.05$
4. $n = 45.034$

Step 3. Final answer:

It will take approximately 45 years for the population to triple in size.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/properties-of-logarithmic-functions/solving-general-problems-with-logarithms-and-exponents/

Simplifying Expressions of the Form $log_a a^x$ and $a(log_a x)$

The expressions $log_a a^x$ and $alog_a x$ can be simplified to x, a shortcut in complex equations.

KEY POINTS

- Because $log_a a = 1$ and $log_a b^x = xlog_a b$, the formula for the logarithm of a power says that for any number x, $log_a a^x = xlog_a a = x$.
- Because $log_a x$ and $log_x a$ are inverse values, $alog_a(x) = x$.
- Simplifying complex-looking equations can greatly facilitate the solving of longer problems.

Recall that the logarithm of a number is defined as the exponent by which another fixed value, the base, must be raised to produce that number. When the base is the same as the number being modified, the solution is 1, or:

$$log_a a = 1$$

Here, $log_a a$ is asking about what power a must be raised to get a; that power is one.

The logarithm of the p-th power of a number is p times the logarithm of the number itself:

$$log_a b^x = xlog_a b$$

Similarly, the logarithm of a p-th root is the logarithm of the number divided by p:

$$log_a b^{1/x} = (1/x)log_a b$$

Because, the formula for the logarithm of a power says that for any number x:

$$log_a a^x = xlog_a a = x$$

Once again, is asking about what power a must be raised to get a; that power is x. In prose, we can say that taking the x-th power of a and then the base-a logarithm gives back x. Conversely, if a positive number a is raised to the power of the log base-a of x, the answer again yields x:

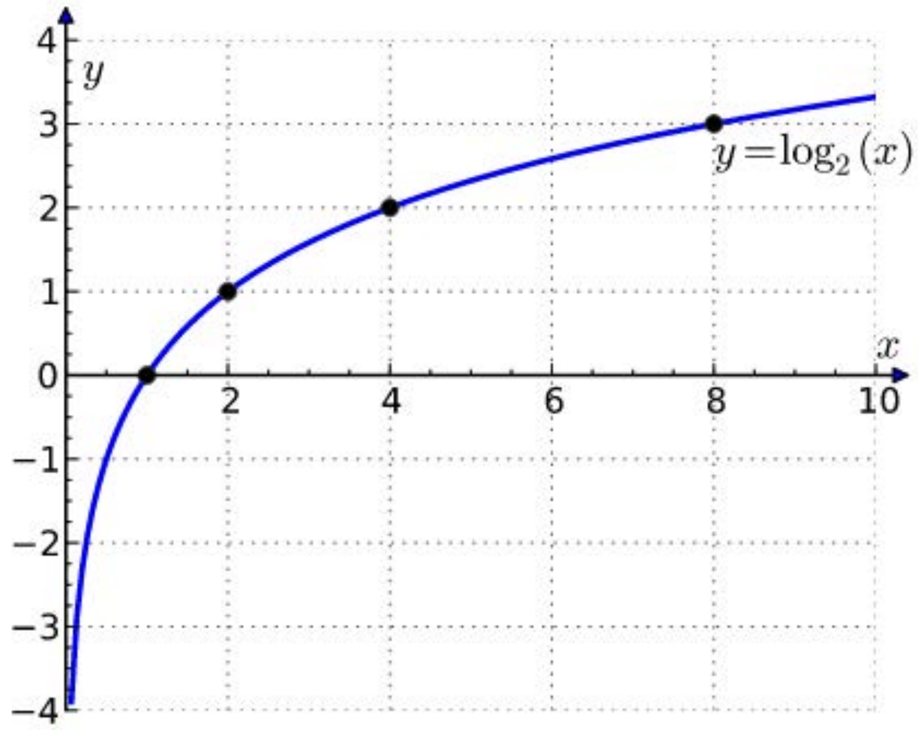
$$alog_a x = x$$

This formula says that first taking the logarithm and then exponentiating gives back x.

Thus, the two possible ways of combining (or composing) logarithms and exponentiation give back the original number. Therefore, the logarithm to base-a is the inverse function of.

$f(x) = a^x$.

Figure 5.28 Graph of Binary Logarithm



The graph of log base 2

Section 5

Growth and Decay; Compound Interest

Population Growth

Interest Compounded Continuously

Limited Growth

Exponential Decay

https://www.boundless.com/algebra/exponents-and-logarithms/growth-and-decay-compound-interest/

Population Growth

Population size can fluctuate positively or negatively, and growth is capable of being modeled by an exponential function.

KEY POINTS

- The formula for population growth is of the same form as that of compound interest. That is, $P(r,t,f) = P_i(1+r)^{\frac{t}{f}}$, where Pi represents initial population, r is the rate of population growth (expressed as a decimal), t is elapsed time, and f is the period over which time population grows by a rate of r.
- Population growth rate is more complex than interest rate; it is dependent on four variables. Population growth rate can be modeled as: $\Delta P = (B - D) + (I - E)$, where ΔP represents change in population, B and D are births and deaths, respectively, and I and E are immigrants and emigrants, respectively.
- Population growth rate can reveal whether a population size is increasing (positive) or decreasing (negative). It can be calculated for two times with the following equation: $PGR = \frac{ln(P(t_2)) - ln(P(t_1))}{(t_2 - t_1)}$, where t2 and t1 represent the two times.

Mathematically, population growth is very similar to the growth of money by compound interest. The equations used to model both are of an **exponential** form.

Population growth can be represented by the following formula:

$$P(r,t,f) = P_i(1+r)^{\frac{t}{f}}$$

Where Pi represents initial population, r is the rate of population growth (expressed as a decimal), t is elapsed time, and f is the period over which time population grows by a rate of r. The ratio of t to f is often simplified into one value representing the number of compounding cycles.

Note that r is more complex in the case of population growth than in interest. In interest, a rate is simply a number specified by a bank. In population growth, it is determined by births (B), deaths (D), **immigrants** (I), and **emigrants** (E):

$$\Delta P = (B - D) + (I - E)$$

The formula is split into natural growth (B-D) and mechanical growth (I-E).

Whereas interest will always increase, population size can fluctuate from growth to decline, and back again. As such, another variable is

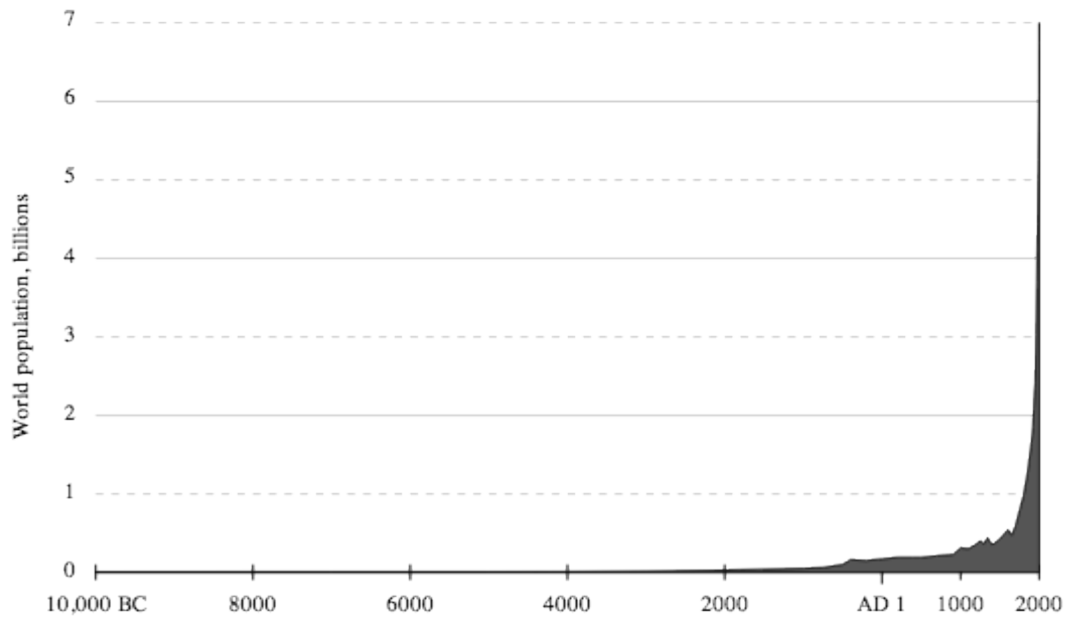
important when studying population demographics and dynamics, Population Growth Rate (PGR).

PGR indicates the rate of change in population over a certain span of time (t_2-t_1). It can be determined from the formula:

$$PGR = \frac{ln(P(t_2)) - ln(P(t_1))}{(t_2)}$$

Multiplying PGR by 100 affords percentage growth, relative to the population at the beginning of the time period.

Figure 5.29 Population Curve



The population of the world has grown at an increasingly staggering rate in recent centuries.

A positive growth rate indicates an increasing population size, while a negative growth rate is characteristic of a decreasing population. A growth rate of 0 means stagnation in population size.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/growth-and-decay-compound-interest/population-growth/

Boundless is an openly licensed educational resource

Interest Compounded Continuously

Compound interest is accrued when interest is earned not only on principal, but on previously accrued interest: it is interest on interest.

KEY POINTS

- Interest is, generally, a fee charged for the borrowing of money. The amount of interest accrued depends on the principal (amount borrowed), the interest rate (a percentage of the principal), period (amount of time between interest payments) and time elapsed.
- The form of the equation for compound interest is exponential, and thus such interest is accrued much faster than the linear simple interest.
- The formula for compound interest is $M = p(1 + r)^{\frac{t}{f}}$ where M represents the total value (including principal), p represents principal, r is interest rate (expressed as a decimal), t is time elapsed, and f is the length of time between payments. To calculate interest alone, simply subtract the principal from M.

Fundamentally, **compound interest** is a specific case of exponential functions found very commonly in everyday life.

Interest is, generally, a fee charged for the borrowing of money. The amount of interest accrued depends on the principal (amount borrowed), the interest rate (a percentage of the principal), period (amount of time between interest payments) and time elapsed.

Simple interest is accrued linearly based on the formula:

$$I = p \cdot r \cdot \frac{t}{f}$$

Where I represents interest, p is principal, r is interest rate (expressed as a decimal), t is time elapsed, and f is the time elapsed per interest payment. The ratio of t to f is often simplified to the number of interest payments.

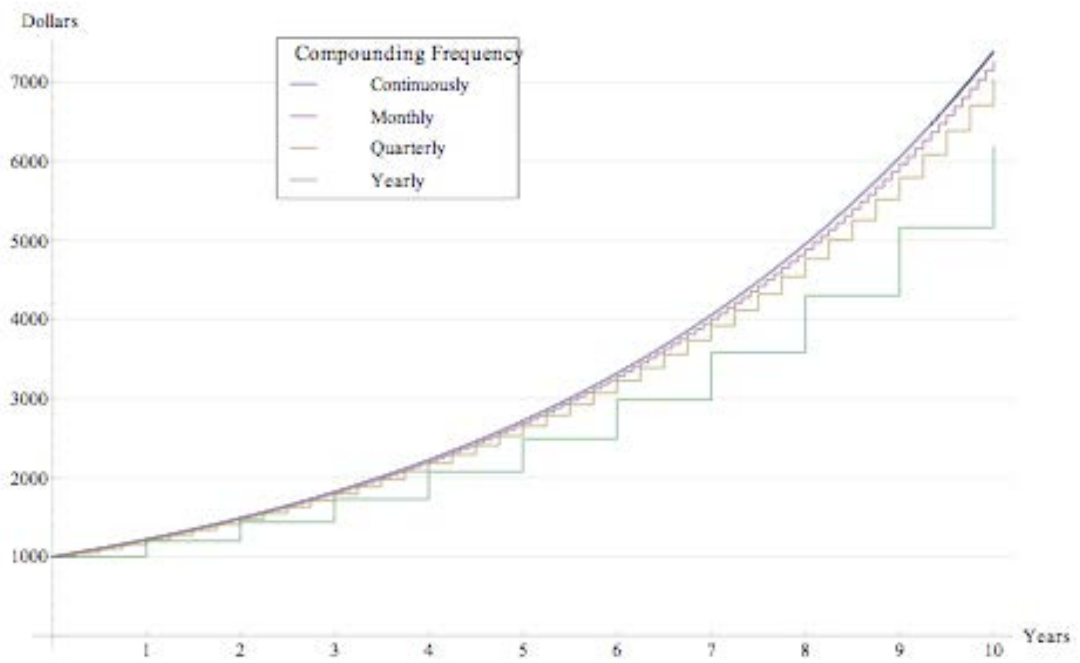
Simple interest is useful in some applications, such as a home equity line. However, there are many instances in which interest will accrue further interest. Consider a bank savings account opened in the amount of $100, which accrues interest at a rate of 5% per year. If left untouched, after one year, the value of the account will be $105. A year later, it will not be $110, as would be suggested by the simple interest formula. Rather, it would be slightly higher due to interest accrued on not only the initial principal ($100), but the interest previously accrued ($5). This is called compound interest.

Compound interest is not linear, but exponential in form. The equation representing investment value as a function of principal, interest rate, period and time is:

$$M = p(1 + r)^{\frac{t}{f}}$$

Where M represents the total value (including principal), p represents principal, r is interest rate (expressed as a decimal), t is time elapsed, and f is the length of time between payments. To calculate interest alone, simply subtract the principal from M.

Figure 5.30 Compound Interest At Varying Frequencies



Starting with a principal of $1000, interest rises exponentially. Notice also that as time passes, a gap forms between the lines as less frequently-compounding methods increase at a lesser rate than more frequently-compounding methods.

EXAMPLE

Consider the aforementioned case of the bank account opened in the amount of $100, which accrues interest at 5% per year. If we leave the account untouched for 50 years, we can calculate the compound interest as such:

$$M = 100(1 + 0.05)^{50}$$

Note that 50 was inserted as the ratio of 50 years and 1 year per interest payment.

$$M = 1146.74$$

Subtracting the $100 principal, we find that interest accrued is in the amount of $1046.74.

Had interest been accrued linearly (in "simple" form), the final value would have been $350, of which $250 would be interest. Thus, it is easy to see how growing interest on interest can make a huge difference in the long term.

Source: https://www.boundless.com/algebra/exponents-and-logarithms/growth-and-decay-compound-interest/interest-compounded-continuously/

Limited Growth

Exponential growth may dampen approaching a certain value, modeled with the logistic growth model: $P(t) = \frac{c}{1 + a \cdot e^{-bt}}$.

KEY POINTS

- Exponential growth may exist within known parameters, but such a functionality may not continue indefinitely.
- If a natural maximum is conceivable, the logistic growth model can be used to represent growth. It has the form: $P(t) = \frac{c}{1 + a \cdot e^{-bt}}$ where P represents population, c is the carrying capacity, b is the population growth rate, t is time, and a is the difference between carrying capacity and initial population.
- An example of natural dampening in growth is the population of humans on planet Earth. The population may be growing exponentially at the moment, but eventually, scarcity of resources will curb our growth as we reach our carrying capacity.

Exponential models of growth and decay can be used to **interpolate** and **extrapolate**. For example, given the population of a country in 1950, 1975, 2000, and at present, one could use the graph of an exponential curve to estimate the rate of population growth in that period, and population at any time between 1950 and the present. This is an example of interpolation.

It is also possible to extrapolate, or make predictions beyond the scope of data. For example, using the same exponential trendline as mentioned above, one could estimate population 10, 20, or 50 years from now.

There are concerns with the interpolative and extrapolative properties of any trendline, but, in addition, there may be other factors that altogether change the form of population growth.

Consider a farm upon which a population of sheep are kept in a constant, comfortable climate in a fully enclosed field. Assume the entire system is closed from gains and losses, but for a flow of a stream of clean drinking water through the field.

If left indefinitely, the population of sheep would perfectly fit an exponential model to a certain point. However, eventually, the grass would act as a factor to limit growth. If the amount of grass available to the sheep and its rate of replenishing are constant, eventually the population of sheep will grow to a tipping point at which the grass can no longer feed the sheep. The death rate of sheep will increase as some starve, and thus the model of population growth among sheep will change form.

Logistic Growth Model

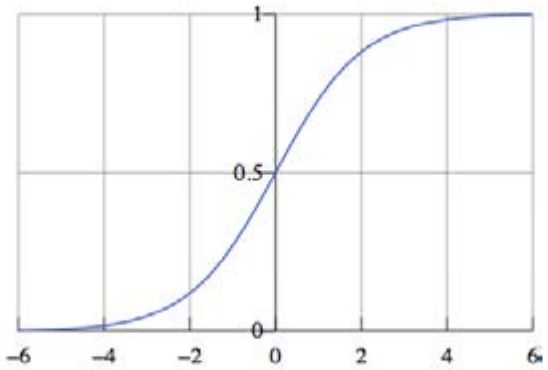
To account for limitations in growth, the logistic growth model can be used:

$$P(t) = \frac{c}{1 + a \cdot e^{-bt}}$$

where P represents population, c is the carrying capacity (maximum the population approaches as time approaches infinity), b is the population growth rate, t is time, and a is the difference between carrying capacity and initial population.

Graphically, the **logistic function** resembles an exponential function followed by a logarithmic function that approaches a horizontal asymptote. From the left, it grows rapidly, but that growth is dampened as it approaches a maximum (*Figure 5.31*).

Figure 5.31 Standard Logistic Function



The graph of the logistic function begins with exponential growth from the left, but that growth is dampened as it approaches a horizontal asymptote to the right.

Exponential Decay

Just as a variable can exponentially increase as a function of another, it is possible for a variable to exponentially decrease.

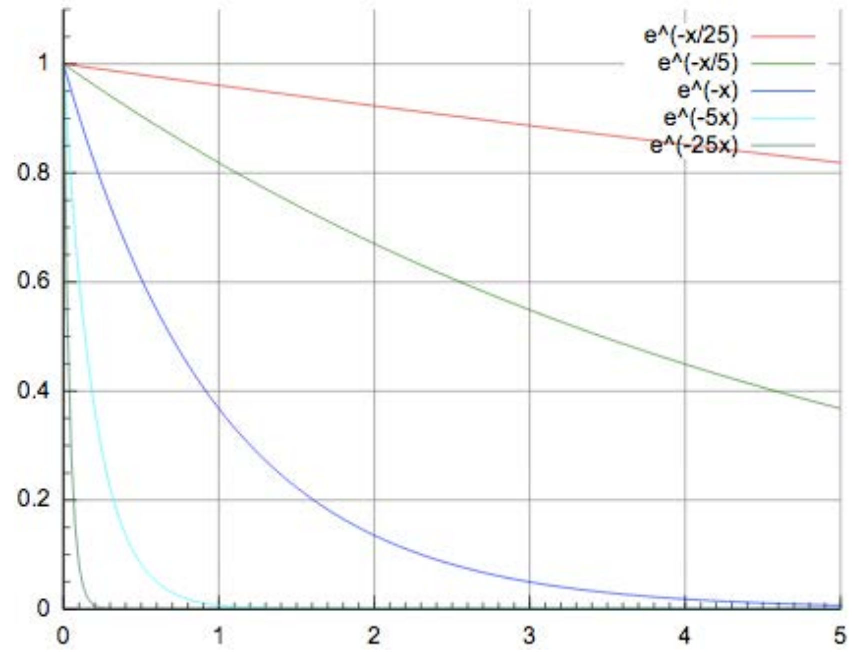
KEY POINTS

- Exponential decrease can be modeled as: $N(t) = N_0 e^{-\lambda t}$ where N is the quantity, N_0 is the initial quantity, λ is the decay constant, and t is time.
- Oftentimes, half-life is used to describe the amount of time required for half of a sample to decay. It can be defined mathematically as: $t_{1/2} = \frac{ln(2)}{\lambda}$ where t1/2 is half-life.
- Half-life can be inserted into the exponential decay model as such: $N(t) = N_0(\frac{1}{2})^{t/t_{1/2}}$ Notice how the exponential changes, but the form of the function will remain.

Just as growth of one variable as a function of another can be exponential in form, so can decline.

Consider the decrease of a population that occurs at a rate proportional to its value. This rate may be constant, but because the population is continually decreasing, the overall decline becomes less and less steep.

Figure 5.32 Exponential Decay



The exponential decline of five different functions is depicted. Note that with higher coefficients of x (λ), the drop is steeper. With lower λ, the drop is more gradual.

Exponential rate of change can be modeled algebraically by the following formula:

$$N(t) = N_0 e^{-\lambda t}$$

where N is the quantity, N_0 is the initial quantity, λ is the decay constant, and t is time. The decay constant is indeed a constant, but the form of the equation (using it as an exponent for e) results in an ever-changing rate of decline. (*Figure 5.32*)

An example of exponential decay is the time-dependent decline in population of a sample of a radioactive **isotope**. Given a sample of

carbon in an ancient, preserved piece of flesh, the age of the sample can be determined based on the percentage of radioactive carbon-13 remaining (1.1% of carbon is C-13; it decays to carbon-12, which represents approximately all the remaining carbon).

C-13 has a **half-life** of 5700 years—that is, in 5700 years, half of a sample of C-13 will have converted to C-12.

Half-life can be mathematically defined as:

$$t_{1/2} = \frac{ln(2)}{\lambda}$$

It can also be conveniently inserted into the exponential decay formula as follows:

$$N(t) = N_0(\frac{1}{2})^{t/t_{1/2}}$$

Thus, if a sample is found to contain 0.55% of its carbon as C-13 (exactly half of the usual 1.1%), it can be calculated that the sample has undergone exactly one half-life, and is thus 5,700 years old.

Chapter 6

Systems of Equations and Matrices

https://www.boundless.com/algebra/systems-of-equations-and-matrices/

Systems of Equations in Two Variables

https://www.boundless.com/algebra/systems-of-equations-and-matrices/systems-of-equations-in-two-variables/

Solving Systems Graphically

The graphical method is a simple way to solve a system of equations by looking for the intersecting point or points of the equations.

KEY POINTS

- The graphical method is a great way to solve a system of equations, and also to check your work if you are solving the system using elimination or substitution.
- There are many ways to write any equation, but if you are going to solve the system graphically, it is helpful to first isolate the y term on one side of the equation/s.
- You can solve the system by locating the intersections between the different equations in the system. It is possible to have more than one answer that satisfies all equations in a system.

A **system of equations** (also known as simultaneous equations) is a set of equations with multiple variables, often solved with a particular specification of the values of all variables that simultaneously satisfies all of the equations. The most common ways to solve a system of equations are:

- The elimination method
- The substitution method
- **The graphical method**

In this atom we will address the graphical method.

Some systems have only one set of correct answers, while others have multiple sets that will satisfy all equations. Shown graphically (*Figure 6.1*), a set of equations solved with only one set of answers will have only have one point of intersection. A system with two sets that will satisfy both equations has two points of intersection (*Figure 6.2*).

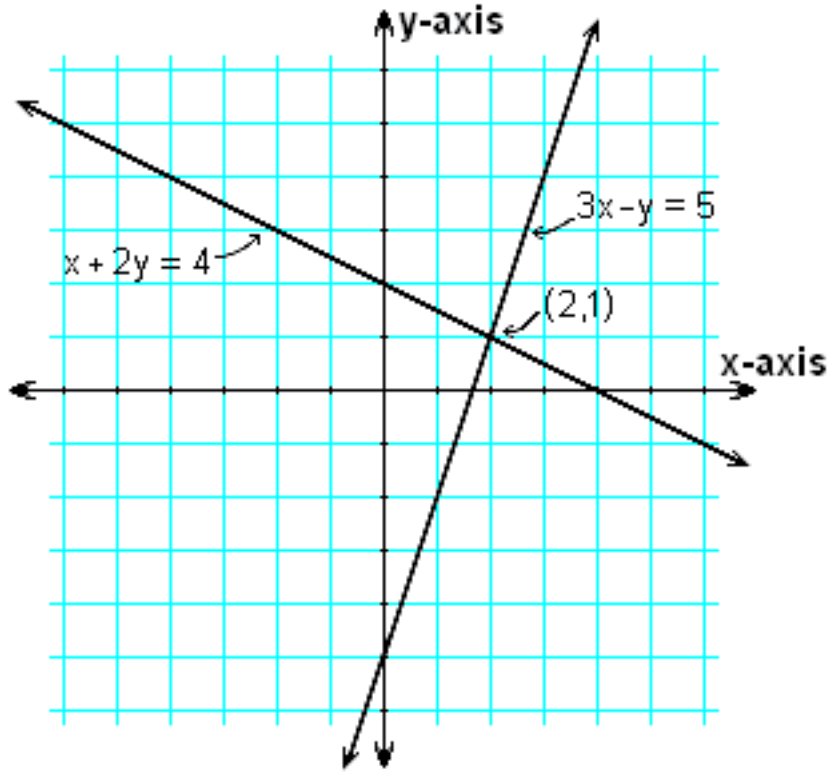


Figure 6.1 System of linear equations with two variables

This graph shows a system of equations with two variables with only one set of answers.

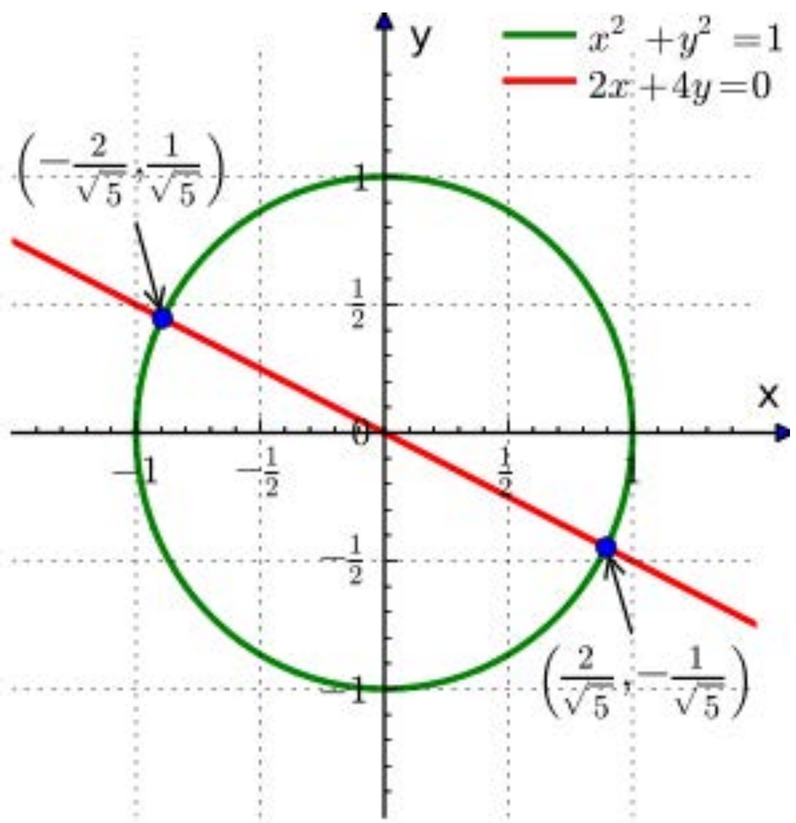


Figure 6.2 System of Equations with multiple answers

This is an example of a system of equations shown graphically that has two sets of answers that will satisfy both equations in the system.

Before successfully solving a system graphically, one must be comfortable with the system of equations represented in a general format, such as written in the following manner: $Ax + By = C$. At first, this may look confusing to try to graph, but keep in mind that converting this equation into the slope-intercept form is relatively easy, and will look like this: $y = mx + b$. Where: m = slope; b = the y intercept. The best way to convert your equation form that is easier to graph is by first isolating the y to yield $By = -Ax + C$, then dividing the right side by B to yield

$y = (-Ax + C)/B$

$y = (-A/B)x + (C/B)$.

The (-A/B) is now the slope, m and the (C/B) is the y-intercept, b. If you have a graphing calculator, you can also use that to represent the equations graphically, but it is useful to know how to represent such equations formulaically on your own.

An ordered pair is a way of writing the correct values for the system in a manner typically associated with graphs (x,y). Once you have figured out how to represent your system of equations graphically, finding the correct ordered pair to satisfy the system is easy: you simply find the intersections between the graphs.

The Substitution Method

The substitution method is a way of solving a system of equations by expressing the equations in terms of only one variable.

KEY POINTS

- A system of equations is a set of equations that can be solved using a particular set of values.
- The substitution method works by expressing one of the variable in terms of another, then substituting it back into the original equation thus simplifying it.
- It is very important to check your work once you have found a set of values for the variables. Do this by substituting the values you found back into the original equations.
- The answer to the system of equations can be written as an ordered pair (x,y).

A system of equations is a set of equations containing multiple variables. Their solution is a set of variables that satisfy all equations in the system. Although there are many methods to solving a **system of equations**, the three simplest are:

- The substitution method
- The elimination method
- The graphical method

In this atom we will address the substitution method.

The substitution method is a way to simplify a system of equations by expressing one variable in terms of another, thus removing one variable from an equation. When all variables in an equation are replaced such that there is only one variable (in different terms), the equation becomes solvable. This method is easier to explain by visually.

For this example, let's take the following system of equations:

$$x - y = -1$$

$$x + 2y = -4$$

The next step is to chose one variable to express in terms of the other. In this case, we will express x in terms of y:

$$x = y - 1$$

We can now substitute our new definition of x (in this case y-1) into the second equation:

$$(y - 1) + 2y = -4$$

Note that now this equation only has one variable (y). We can then simplify this equation to solve for y:

$3y = -3$

$y = -1$

Now that we know the value of y, we can use it to find the value of the other variable, x:

$x = -1 - 1$

$x = -2$

Thus, the solution to the system is: (-2, -1). On an x-y coordinate plane, this is the point where the two functions intersect. Graphically, one can confirm this to be true, as in *Figure 6.3*.

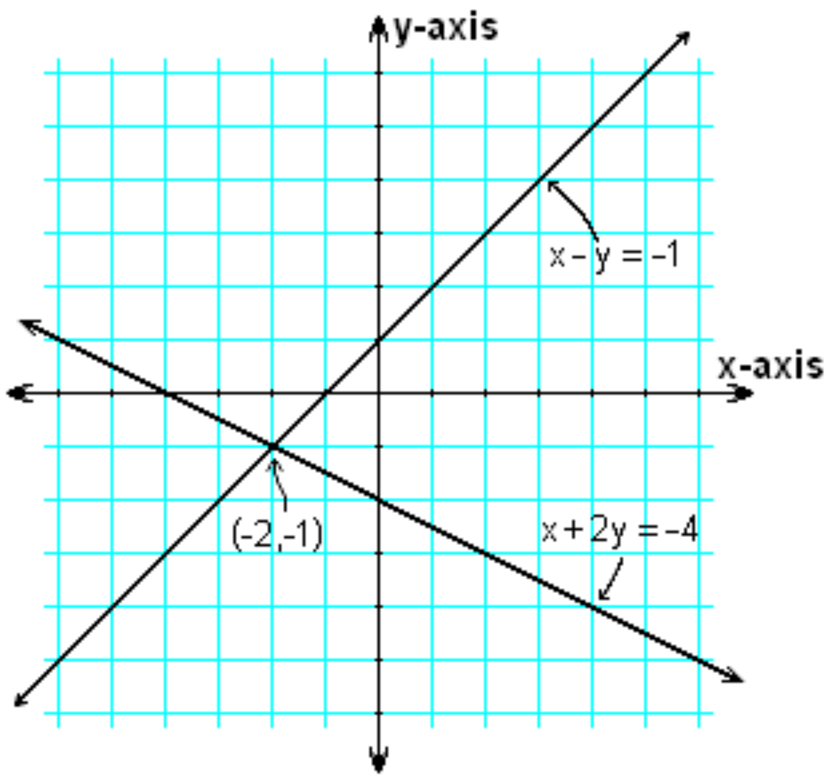


Figure 6.3 System of Equations with Two Variables

This figure is a graphical representation of the example equations. You can use this to check your answers, and see that the ordered pair you come up with matches where the lines intersect with one another.

EXAMPLE

Start with this system of equations:

$x - y = -1$

$x + 2y = -4$

Pick a variable to write in terms of the other:

$x = y - 1$

Substitute that back into the original equation:

$(y - 1) + 2y = -4$

Isolate the variable:

$y + 2y = -4 + 1$

$-3y = -3$

$y = -1$

Use this value to solve for the other variable:

$x = y - 1$

$x = (-1) - 1$

$x = -2$

Now that you have your answer (-2,-1), check your answer by plugging these values back into one of the original equations:

$x + 2y = -4$

$(-2) + 2(-1) = -4$

$-2 - 2 = -4$

$-4 = -4$

It works! Good job!

The Elimination Method

The elimination method is used to eliminate a variable in order to more simply solve for the remaining variable(s).

KEY POINTS

- In order to easily solve the system of equations, it helps to first set the equations up in similar way. For example, $x+y=-1$ and $2y+x=-4$ should be written: $x+ y=-1$ and $x+2y=-4$.
- Once, the values for the remaining variables have been found successfully, then go back and plug that result into one of the original equation and find the correct value for the other variable.
- Always check the work. This is done by plugging both values into one or both of the original equations.

A **system of equations**, also known as simultaneous equations, are a set of at least two equations containing multiple variables. The system can be solved using a specific set of values corresponding to the variables. This section will focus on systems of equations with just two variables. It is important to first learn how to solve simple systems in order to understand the basic concept before moving on to more complex systems of equations with more variables. The values of the variables are often written in the following notation: (x,y). The most basic methods for solving a system of equations are:

- The graphical method
- The substitution method
- and the **elimination method**

In this section, the elimination method will be explored.

The elimination method, also known as elimination by addition, is a way to eliminate one of the variables to more simply evaluate the remaining variable. Once, the values for the remaining variables have been found successfully, then go back and plug that result into one of the original equation and find the correct value for the other variable.

This elimination method can be demonstrated by using a simple example:

Using two equations, both with two variables; 2x+y=8 and y+x=6, the first thing that needs to be done is to write the variables in the same manner, so that the similar variables line up, like so: 2x+y=8 x +y=6.In this example, one variable, y, can easily eliminate using the elimination method, just subtract the bottom equation from the top equation. The remaining variable are: 2x+y=8 minusx+y=6 equals x=2.Then go back to one of the original equations and substitute the value we found for x. It is easiest to pick the simplest equation, but either equation will work: x+y=6 --> 2+y=6 y=4. I has been found,

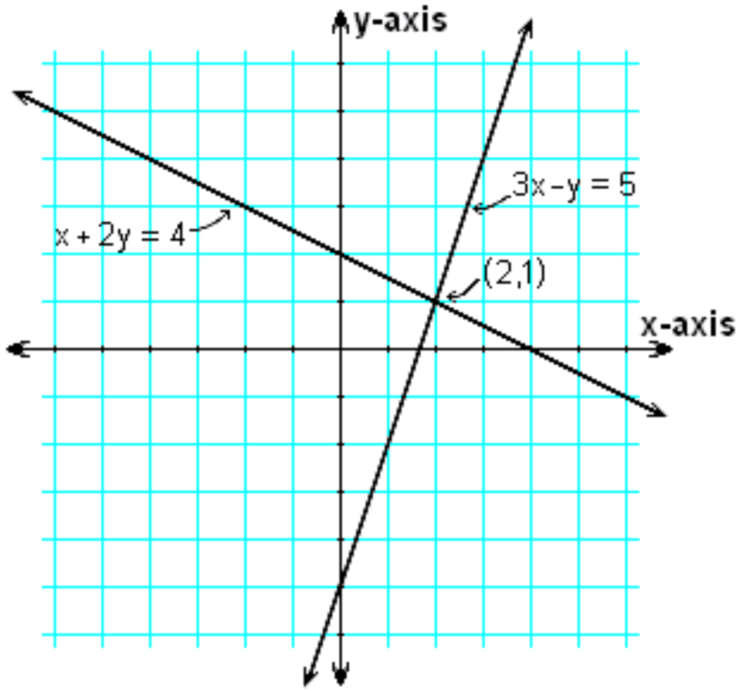


Figure 6.4 System of equations with two variables

This image demonstrates that the system of equations has a specific set of values which will solve BOTH equations

using the method of elimination, that (x,y) is equal to (2,4). It is also easy to go back and check the answer by plugging both of these values in for their respective variables into one of the equations: 2x +y=2(2)+4=4+4=8.

Although, this is a very simplified example, make sure to follow these steps, and it will be easy to solve any system of equations with two variables. Now, imagine that there is a more difficult problem, such as: x+2y=4 and 3x-y=5. This system of equations is shown in *Figure 6.4*. The best way to start every problem is to simplify the equation. Set the equation up so that one of the variables will easily go away. Start by making one of the variables consistent in both equations. This can be done by multiplying the second equation by

2. Don't forget, the same thing has to be done to both sides of the equation: 2[3x-y]=2 becomes 6x-2y=10.After the equations are placed on top of each other, the equation will be: x+2y=46x-2y=10. Add the equations together the same way as in the first example.

Also, it is crucially important to check the work, especially when it is simple to check for a system of equations.

EXAMPLE

Given the following system of equations: x+2y=4 and 3x-y=5

The answer should be written in the following notation:
(x,y) x+2y=4 and 3x-y=5

First, set the system up in a way that makes it easier to manipulate:
x+2y=4 3x-y=5

Next, choose a variable to eliminate. In this example, eliminate y, therefore:
x+2y=4 x+2y=4 [3x-y]=[5]-----> 6x-2y=10 7x =143.

Solve for the remaining variable:
7x=14 x=14/7 x=24

Plug into original equation:
x+2y=4(2)+2y=4 2y=2 y=2/2 y=15

We have now found that (x,y) = (2,1). Finally, plug these values into one of the equations to make sure everything checks out!

3x-y=53(2)-(1)=5 6-1=5 5=5. It works! Good job!

Source: https://www.boundless.com/algebra/systems-of-equations-and-matrices/systems-of-equations-in-two-variables/the-elimination-method/

Applications of Systems of Equations

Systems of equations can be used to solve many real-life problems in which multiple constraints are used on the same variables.

KEY POINTS

- If you have a problem that includes multiple variables, you can solve it by creating a system of equations.
- The first step to solving a multivariate problem is to identify and label the variables.
- Once variables are defined, determine the relationships between them and write them as equations.

A system of equations, also known as simultaneous equations, is a set of equations that have multiple variables. The answer to a system of equations is a set of values that satisfies all equations in the system. Systems of equations can have multiple sets of answers that are correct. Solutions to a system of equations are often written as **ordered pairs**, (x,y). There are many ways of solving a system of equations, including the elimination, substitution, and graphical methods.

There are many applications of systems of equations. Whenever you have a problem that has multiple variables, setting up a system of equations is often the best method for solving. The steps you need to take in order to do that are: (1) identify the variables in the problem, (2) name the variables, and (3) set up the equations and solve for each variable.

Example 1

Emily is hosting a major after-school party. The principal has imposed two restrictions. First, the total number of people attending (teachers and students combined) must be 56. Second, there must be one teacher for every seven students. How many students and how many teachers are invited to the party?

First, we need to identify our variables. In this case, our variables are teachers and students. Now we need to name these variables: number of teachers will be T, and number of students will be S.

Now we need to set up our equations. There is a **constraint** limiting the total number of people in attendance to 56, so:

$$T + S = 56$$

For every seven students, there must be one teacher, so:

$$7S = T$$

Now we a system of equations that can be solved by substitution, elimination, or graphically. The solution to the system is S=49 and T=7.

Example 2

A group of 75 students and teachers are in a field, picking sweet potatoes for the needy. Kasey picks three times as many sweet potatoes as Davis—and then, on the way back to the car, she picks up five more sweet potatoes than that! Looking at her newly increased pile, Davis remarks "Wow, you've got 29 more potatoes than me!" How many sweet potatoes did Kasey and Davis each pick?

To solve, we first define our variables. The number of sweet potatoes that Kasey picks is K, and the number of sweet potatoes that Davis picks is D.

Now we can write equations based on the situation:

$K - 5 = 3D$

$D + 29 = K$

From here, substitution, elimination or graphing will reveal that K is 41 and D is 12.

It is important that you always check your answers. A good way to check solutions to a system of equations is to look at the functions graphically and then see where the graphs intersect (*Figure 6.5*).

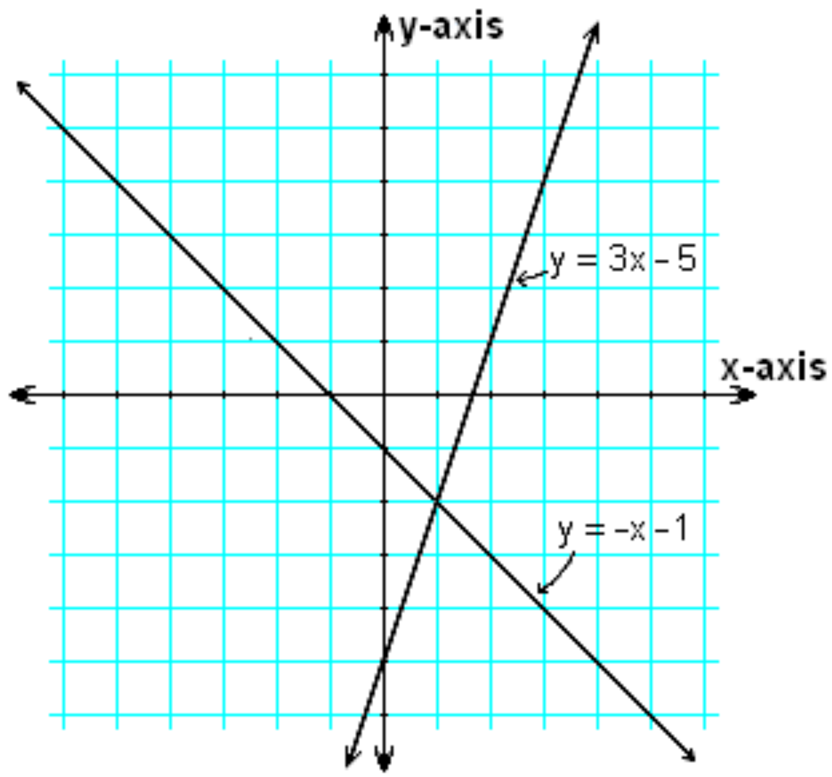


Figure 6.5 Graphical representation of a system of equations

Once you have created your system of equations, one way to solve it is by showing them graphically, and then finding the intersecting point or points.

Source: https://www.boundless.com/algebra/systems-of-equations-and-matrices/systems-of-equations-in-two-variables/applications-of-systems-of-equations/

Section 2

Systems of Equations in Three Variables

Solving Systems of Equations in Three Variables

Applications and Mathematical Models

https://www.boundless.com/algebra/systems-of-equations-and-matrices/systems-of-equations-in-three-variables/

Solving Systems of Equations in Three Variables

A system of equations in three variables involves two or more equations, each of which involves between one and three variables.

KEY POINTS

- A system of equations may have no solutions, one unique solution, or infinitely many solutions.
- The substitution method involves solving for one of the variables in one of the equations, and plugging that into the rest of the equations to reduce the system. Rinse and repeat until there is a single equation left, and then using this go backwards to solve the previous equations.
- The graphical method involves graphing all of the equations and finding points, lines or planes where all of the equations intersect at once, such points, lines or planes are the solutions.
- The elimination method involves adding or subtracting multiples of one equation from the other equations, eliminating variables from each of the equations until one variable is left in each equation (if there is a unique solution).

In mathematics, simultaneous equations are a set of equations containing multiple variables. This set is often referred to as a **system of equations**. A solution to a system of equations is a particular specification of the values of all variables that simultaneously satisfies all of the equations. The elementary methods to solve simple systems of equations include graphical method, the matrix method, the substitution method, or the elimination method.

Among the systems of equations, the systems of linear equations are especially important. There were the original object of study of linear algebra. Many algorithms have been devised to solve them, which allow to solve huge systems (up to millions of variables).

For a system of equations in three variables, you can have one or more equations, each of which may contain one or more of the three variables, usually x, y and z.

Finding Solutions

Sometimes not all variables can be solved for, and so an answer for at least one variable must be expressed in terms of other variables and so the set of all solutions is infinite. This is typical for the case where the system has fewer equations than variables. If the number of equations is the same as the number of variables, then probably, but not necessarily, the system is exactly solvable in the sense that

the set of its solutions is finite. For a system of linear equations in this case, there is exactly one solution, for other systems to have several solutions is also typical. A consistent system is a system of equations with at least one solution. Sometimes a system is inconsistent, or has no solution. Having no solutions is typical for the case where the system has more equations than variables.

Graphical Method

The graphical method involves plotting the planes that are formed by a system of equations in three variables and finding the intersection point, line, or plane that solves the system. A single point is a unique solution, whereas intersecting lines or planes mean an infinite number of solutions. If all the planes don't intersect at the same spot, there are no solutions. An example of this is given in *Figure 6.6*.

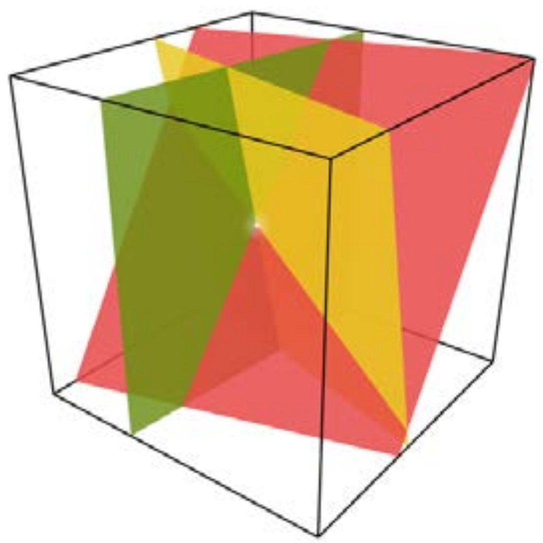


Figure 6.6 System of Linear Equations

This images shows a system of three equations in three variables. The intersecting point is the unique solution to this system.

Substitution Method

The substitution method involves finding an equation that can be written with a single variable as the subject, in which the left-hand side variable does not occur in the right-hand side of the expression. Next, substitute that expression where that variable appears in the other equations, thereby obtaining a smaller system with fewer variables. After that smaller system has been solved, whether by further application of the substitution method or by other methods, substitute the solutions found for the variables in the above right-hand side expression.

Using the simple example above, one can perform the substitution method.

Taking the first equation $3x + 2y - z = 6$, since the coefficient of z is already 1, let's solve for z to get $z = 3x + 2y - 6$. Plug this into the other two equations to get the new smaller system:

$$\begin{cases} -2x + 2y + 3x + 2y - 6 = 3 \\ x + y + 3x + 2y - 6 = 4 \end{cases}$$

Which simplifies to:

$$\begin{cases} x + 4y = 9 \\ 4x + 3y = 10 \end{cases}.$$

Now solving for x in the first equation, one gets $x = 9 - 4y$. Plug this into last equation in the system to get $4(9 - 4y) + 3y = 10$, which simplifies to $y = 2$. Now having the value of y, work back up the equation. Plug y = 2 into the equation $x = 9 - 4y$ to get x=1. Working up again, plug y = 2 and x = 1 into the first substituted equation, $z = 3x + 2y - 6$, which simplifies to z = 1.

Elimination Method

Elimination by judicious multiplication is the other commonly used method to solve simultaneous linear equations. It uses the general principles that each side of an equation still equals the other when both sides are multiplied (or divided) by the same quantity, or when the same quantity is added (or subtracted) from both sides. As the equations grow simpler through the elimination of some variables, a variable will eventually appear in fully solvable form, and this value can then be "back-substituted" into previously derived equations by plugging this value in for the variable. Typically, each "back-substitution" can then allow another variable in the system to be solved.

Looking at the example:

$$\begin{cases} x + y + z = 2 \\ x - y + 3z = 4 \\ 2x + 2y + z = 3 \end{cases}$$

Using the elimination method, first subtract the first equation from the second equation to get $x - y + 3z - (x + y + z) = 4 - 2$, which simplifies to $-2y + 2z = 2$, giving us the system of equations:

$$\begin{cases} x + y + z = 2 \\ -2y + 2z = 2 \\ 2x + 2y + z = 3 \end{cases}$$

Now subtract two times the first equation from the third equation to get $2x + 2y + z - 2(x + y + z) = 3 - 2(2)$, which simplifies to $z = 1$. Doing this shows the new system:

$$\begin{cases} x + y + z = 2 \\ -2y + 2z = 2 \\ z = 1 \end{cases}$$

Next, subtract two times the third equation from the second equation to get $-2y + 2z - 2z = 2 - 2$, simplifying to $y = 0$. Doing this shows the new system:

$$\begin{cases} x + y + z = 2 \\ y = 0 \\ z = 1 \end{cases}$$

Finally, subtract the third and second equation from the first equation to get $x + y + z - y - z = 2 - 0 - 1$, giving $x = 1$ and the final, solved, system that is:

$$\begin{cases} x = 1 \\ y = 0 \\ z = 1 \end{cases}$$

Unsolvable Systems

If there is a system of three variables with only two equations, one will end up with either an infinite number of solutions or no solutions.

$$\begin{cases} x + y + z = 2 \\ -x - y + z = 4 \end{cases}$$

Which through any of the methods will simplify to:

$$\begin{cases} x + y = -1 \\ z = 3 \end{cases}$$

Which has an infinite number of solutions so long as x + y = -1.

An example of a system with no solutions is:

$$\begin{cases} x + y + z = 2 \\ 2x + 2y + 2z = 6 \end{cases}$$

This can be seen by subtracting two times the first equation from the second equation, which simplifies to 0=2, which is impossible, hence there are no solutions.

These situations can arise in any number of equations and variables.

Applications and Mathematical Models

Systems of equations are problems that have multiple unknowns and multiple observations, and can be used in many practical applications.

KEY POINTS

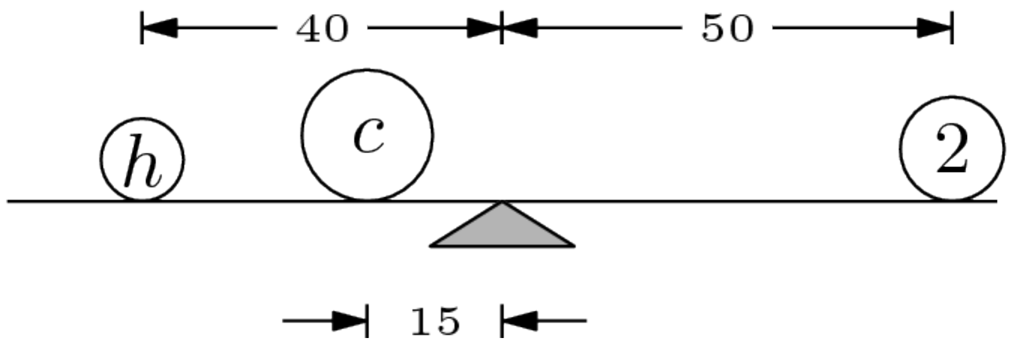
- When you have multiple unknown quantities with multiple observations on these quantities and their interactions with each other, then the problem can usually be naturally described with a system of equations.
- When a system of equations is laid out, all of the equations need to be satisfied in order for there to be a solution. Sometimes there are no solutions; other times there are infinitely many solutions.
- There are numerous applications for systems of equations, such as Physics problems that involve multiple objects with multiple observations, or multiple forces that all need to be balanced.

Systems of **linear equations** are common in science and mathematics, including Physics, Chemistry and maximization/minimization and **constraint** problems.A system of equations is a way to evaluate multiple unknown quantities. You will need observations of these quantities in order to properly solve for the unknowns. The simplest and most studied method is a system of linear equations, and we will use this type in our examples.

Physics Example

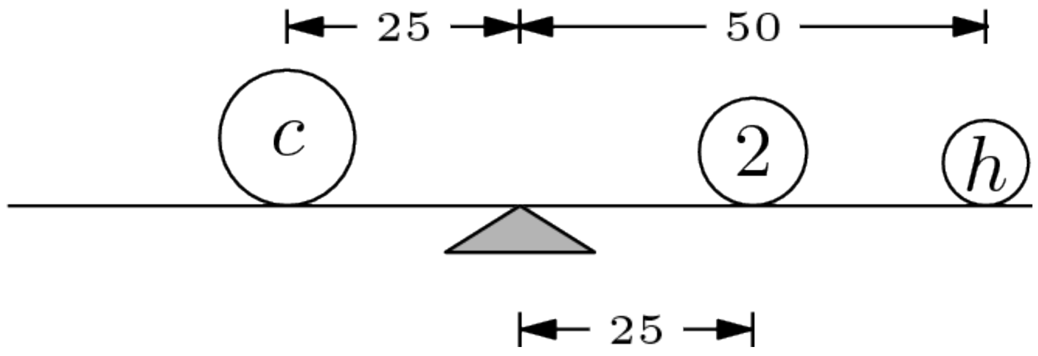
Systems of equations are best explained through examples, and for practicality sake, let's start with a physics example. Let's say you have three balls, but you are only told the mass of one of them. Your unknowns are the other two balls, c and h. The one known mass is 2 kg. You take two observations with the three balls balanced on a bar over a fulcrum, *Figure 6.7* and *Figure 6.8*. You therefore know the sum of the moments on the left equal the sum of the moment on the right.*The moment of an object is its mass multiplied by its distance from the balance point. The two balances give this system of two

Figure 6.7 Physics Example: First Observation



Here is the first observation of three balls, two with unknown weight, which are balanced on a bar in a given configuration.

Figure 6.8 Physics Example: Second Observation



Here is the second observation of three balls, two with unknown weight, which are balanced on a bar in a given configuration.

equations. In this example, the unknowns are the two masses and the observations are the balances.

Using the substitution method on this example, let's use the second equation and solve for c to get $c = 2 + 2h$. Plug this into the first equation to get $40h + 15(2 + 2h) = 100$ which simplifies to $h = 1$, and plugging this into the previous equation gives us $c = 4$.

Chemistry Example

We mix, under controlled conditions, a certain amount of toluene, C_7H_8, and nitric acid HNO_3, to produce trinitrotoluene, $C_7H_5O_6N_3$ and water H_2O (conditions have to be controlled; trinitrotoluene is better known as TNT). In what proportion should those components be mixed? The number of atoms of each element present before the reaction must equal the number present afterward.

$$xC_7H_8 + yHNO_3 \rightarrow zC_7H_5O_6N_3 + wH_2O$$

Applying that principle to the elements C, H, N, and O results in this system.

$$\begin{cases} 7x & = & 7z \\ 8x + y & = & 5z + 2w \\ y & = & 3z \\ 3y & = & 6z + w \end{cases}$$

Let's use substitution. The first line gives us x = z, giving us the reduced system:

$$\begin{cases} 8z + y & = & 5z + 2w \\ y & = & 3z \\ 3y & = & 6z + w \end{cases}$$

The second equation in this system gives us y = 3z, giving us the reduced system:

$$\begin{cases} 8z + 3z & = & 5z + 2w \\ 9z & = & 6z + w \end{cases}$$

Which simplifies to:

$$\begin{cases} 6z & = & 2w \\ 3z & = & w \end{cases}$$

Since these two equations are equivalent, we have an infinite number of solutions. This makes sense seeing that, for instance, if you place a certain amount of toluene and nitric acid to produce the TNT and water, then placing a multiple of that amount will give you a multiple of TNT. So we'll simply choose one solution, and know that there are infinitely many multiples of this one solution. Let's choose z=1. Therefore w=3, and propagating these two values up, we find that y=3 and x=1.

Therefore one atom of toluene and three atoms of nitric acid produce one atom of trinitrotoluene and three atoms of water.

Section 3

Matrices

https://www.boundless.com/algebra/systems-of-equations-and-matrices/matrices/

Matrices and Row-Equivalent Operations

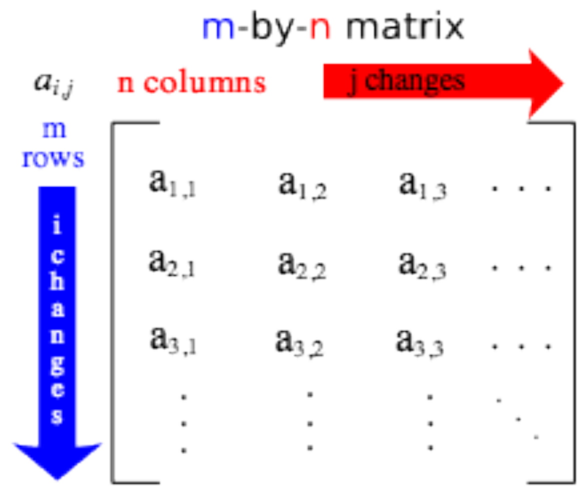
Two matrices are row equivalent if one can be changed to the other by a sequence of elementary row operations.

KEY POINTS

- Because elementary row operations are reversible, row equivalence is an equivalence relation.
- An elementary row operation is any one of the following moves: Swap (swap two rows of a matrix), Scale (multiply a row of a matrix by a nonzero constant), or Pivot (add to one row of a matrix some multiple of another row).
- If the rows of the matrix represent a system of linear equations, then the row space consists of all linear equations that can be deduced algebraically from those in the system.

In linear algebra, two matrices are **row equivalent** if one can be changed to the other by a sequence of elementary row operations. Alternatively, two m × n matrices (*Figure 6.9*) are row equivalent if and only if they have the same row space. The concept is most commonly applied to matrices that represent systems of linear equations, in which case two matrices of the same size are row equivalent if and only if the corresponding homogeneous systems have the same set of solutions, or equivalently the matrices have the same null space. Because elementary row operations are reversible, row equivalence is an equivalence relation. It is commonly denoted by a tilde (~).There is a similar notion of column equivalence, defined by elementary column operations. Two matrices are column equivalent if and only if their transpose matrices are row equivalent. Two rectangular matrices that can be converted into one another allowing both elementary row and column operations are called simply equivalent.

Figure 6.9 A Matrix



Specific elements of a matrix are often denoted by a variable with two subscripts. For instance, a2,1 represents the element at the second row and first column of a matrix A.

An elementary row operation is any one of the following moves:

Swap: Swap two rows of a matrix.

Scale: Multiply a row of a matrix by a nonzero constant.

Pivot: Add to one row of a matrix some multiple of another row.

For example, the following steps show that these two matrices are row equivalent:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 1 \end{pmatrix} \quad B = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}$$

Start with A, keep the second row, and then add the first to the second:

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

Then, keep the first row. Multiply the second row by 3 and then subtract the first row from the second row:

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 3 & 3 & 2 \end{pmatrix}$$

Keep the first row again and then subtract the first row from the second:

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 1 \end{pmatrix}$$

Now, you can see that A = B, which we achieved through a series of elementary row operations.

Row Space

The row space of a matrix is the set of all possible linear combinations of its row vectors. If the rows of the matrix represent a system of linear equations, then the row space consists of all linear equations that can be deduced algebraically from those in the system. Two $m \times n$ matrices are row equivalent if and only if they have the same row space. For example, the matrices are row equivalent, the row space being all vectors of the form. The corresponding systems of homogeneous equations convey the same information. In particular, both of these systems imply every equation of the form.

Gaussian Elimination

Using elementary operations, Gaussian elimination reduces matrices to row echelon form.

KEY POINTS

- Since elementary row operations preserve the row space of the matrix, the row space of the row echelon form is the same as that of the original matrix.
- There are three types of elementary row operations: swap the positions of two rows, multiply a row by a nonzero scalar, and add to one row a scalar multiple of another.
- In practice, one does not usually deal with the systems in terms of equations but instead makes use of the augmented matrix (which is also suitable for computer manipulations).

By means of a finite sequence of elementary row operations, called Gaussian elimination, any matrix (*Figure 6.9*) can be transformed to a row echelon form. Since elementary row operations preserve the row space of the matrix, the row space of the row echelon form is the same as that of the original matrix. The resulting echelon form is not unique. For example, any multiple of a matrix in echelon form is also in echelon form. However, it is the case that every matrix has a unique reduced row echelon form. This means that the nonzero rows of the reduced row echelon form are the unique reduced row echelon generating set for the row space of the original matrix.

Before getting into more detail, there are a couple of key terms that should be mentioned:

- **Augmented matrix**: an augmented matrix is a matrix obtained by appending the columns of two given matrices, usually for the purpose of performing the same elementary row operations on each of the given matrices.
- Upper triangle form: A square matrix is called upper triangular if all the entries below the main diagonal are zero. A triangular matrix is one that is either lower triangular or upper triangular. A matrix that is both upper and lower triangular is a diagonal matrix.
- Elementary row operations: Swap rows, add rows or multiply rows.

Now, the steps of Gaussian Elimination are as follows:

1. Write the augmented matrix for the linear equations.
2. Use elementary row operations on the augmented matrix [A|b] to transform A to upper triangle form. If a zero is on the diagonal, switch the rows until a nonzero is in its place.
3. Use back substitution to find the solution.

Using Gaussian Elimination to solve the following:

$2x + y - z = 8 \quad (L_1)$

$-3x - y + 2z = -11 \quad (L_2)$

$-2x + y + 2z = -3 (L_3)$

Therefore, the Gaussian Elimination algorithm applied to the augmented matrix begins with,

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 8 \\ -3 & -1 & 2 & -11 \\ -2 & 1 & 2 & -3 \end{array}\right]$$

which, at the end of the first part (Gaussian elimination, zeros only under the leading 1) of the algorithm, looks like this:

$$\left[\begin{array}{ccc|c} 1 & 1/3 & -2/3 & 11/3 \\ 0 & 1 & 2/5 & 13/5 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

Therefore, it is in row echelon form. At the end of the algorithm, if the Gauss–Jordan elimination (zeros under and above the leading 1) is applied:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

It is in reduced row echelon form, or row canonical form.

Source: https://www.boundless.com/algebra/systems-of-equations-and-matrices/matrices/gaussian-elimination/

Gauss-Jordan Elimination

Gauss–Jordan elimination is an algorithm for getting matrices in reduced row echelon form using elementary row operations.

KEY POINTS

- Gaussian elimination places zeros below each pivot in the matrix, starting with the top row and working downwards.
- Gauss–Jordan elimination goes a step further by placing zeros above and below each pivot; such matrices are said to be in reduced row echelon form.
- Gauss-Jordan elimination, like Gaussian elimination, is used for inverting matrices and solving systems of linear equations.

In linear algebra, **Gauss–Jordan elimination** is an algorithm for getting matrices in reduced row echelon form using elementary row operations (*Figure 6.10*). It is a variation of Gaussian elimination, which places zeros below each pivot in the matrix, starting with the top row and working downwards. Matrices containing zeros below each pivot are said to be in row echelon form. Gauss–Jordan elimination goes a step further by placing zeros above and below each pivot; such matrices are said to be in reduced row echelon form.

Gauss-Jordan elimination, like Gaussian elimination, is used for inverting matrices and solving systems of linear equations. Both Gauss–Jordan and Gaussian elimination have time complexity of order $O(n^3)$ for an n by n full rank matrix (using Big O Notation), but the order of magnitude of the number of arithmetic operations (there are roughly the same number of additions and multiplications/divisions) used in solving an n by n matrix by Gauss-Jordan elimination is n^3, whereas that for Gaussian elimination is $\frac{2n}{3}$. However, the result of Gauss-Jordan elimination (reduced row echelon form) may be retrieved from the result of Gaussian elimination (row echelon form) in arithmetic operations by proceeding from the last pivot to the first one. Thus the needed number of operations has the same order of magnitude for both eliminations.

Figure 6.10 Matrix in Reduced Row Echelon

$$\begin{bmatrix} 1 & 0 & 0 & 0 & b_1 \\ 0 & 1 & 0 & 0 & b_2 \\ 0 & 0 & 0 & 1 & b_3 \end{bmatrix}$$

A matrix is in reduced row echelon form (also called row canonical form) if it is the result of a Gauss–Jordan elimination.

The steps of Gauss-Jordan elimination are very similar to that of Gaussian elimination, the main difference being that we will work in diagonal form instead of putting the augmented matrix into upper triangle form. In diagonal form, we remove any zeros from the diagonal and add them below and above.

Here are the steps to Gauss-Jordan elimination:

1. Turn the equations into an augmented matrix.
2. Use elementary row operations on matrix [A|b] to transform A into diagonal form. Make sure there are no zeros in the diagonal.
3. Divide the diagonal element and the right-hand element (of b) for that diagonal element's row so that the diagonal element is equal to one.

Turn these equations into an augmented matrix:

$2y + z = 4$

$x + y + 2z = 6$

$2x + y + z = 7$

$$\left[\begin{array}{ccc|c} 0 & 2 & 1 & 4 \\ 1 & 1 & 2 & 6 \\ 2 & 1 & 1 & 7 \end{array}\right]$$

Then, use elementary row operations to transform A into diagonal form:

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 0 & 2 & 1 & 4 \\ 2 & 1 & 1 & 7 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 0 & 2 & 1 & 4 \\ 0 & -1 & -3 & -5 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 3/2 & 4 \\ 0 & 2 & 1 & 4 \\ 0 & 0 & 5/2 & -3 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 11/5 \\ 0 & 2 & 0 & -14/5 \\ 0 & 0 & -5/2 & -3 \end{array}\right]$$

Now that it is in diagonal coefficient form, the final step is to get everything on the diagonal to equal one. Divide each element of the diagonal and the corresponding row element of b by a number which yields a 1 in the diagonal:

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 11/5 \\ 0 & 1 & 0 & 7/5 \\ 0 & 0 & 1 & -6/5 \end{array}\right]$$

$$x = \frac{11}{5}, y = \frac{7}{5}, z = \frac{6}{5}$$

Section 4

Matrix Operations

https://www.boundless.com/algebra/systems-of-equations-and-matrices/matrix-operations/

Addition and Subtraction; Scalar Multiplication

There are a number of operations that can be applied to modify matrices, such as matrix addition, subtraction, and scalar multiplication.

KEY POINTS

- When performing addition, you add each number in the first matrix to the corresponding number in the second matrix.
- When performing subtraction, simply subtract a number in one of the matrices from the corresponding number in the other matrix.
- Addition and subtraction require that the matrices be the same dimensions. Also, you must begin and end with the same dimensions.
- Scalar multiplication of a real Euclidean vector by a positive real number multiplies the magnitude of the vector without changing its direction.

There are a number of operations that can be applied to modify matrices, such as matrix addition and subtraction and **scalar** multiplication. These form the basic techniques to deal with matrices.

Adding and Subtracting Matrices

Adding matrices is very simple. You just add each number in the first matrix to the corresponding number in the second matrix.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \end{pmatrix} = \begin{pmatrix} 11 & 22 & 33 \\ 44 & 55 & 66 \end{pmatrix}$$

For instance, you can take each number that appears in the upper-right-hand corner to create the calculation $3+30=33$. Note that both matrices being added are 2×3, and the resulting matrix is also 2×3. You cannot add two matrices that have different dimensions.

As you might guess, subtracting works much the same way, except that you subtract instead of adding.

$$\begin{pmatrix} 10 & 20 & 30 \\ 40 & 50 & 60 \end{pmatrix} - \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 9 & 18 & 27 \\ 36 & 45 & 54 \end{pmatrix}$$

Once again, note that the resulting matrix has the same dimensions as the originals, and that you cannot subtract two matrices that have different dimensions.

Scalar Multiplication

In an intuitive geometrical context, scalar multiplication of a real Euclidean vector by a positive real number multiplies the magnitude of the vector without changing its direction. What does it

mean to multiply a number by 3? It means you add the number to itself 3 times. Multiplying a matrix by 3 means the same thing; you add the matrix to itself 3 times.

$$3 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{pmatrix}$$

Note what has happened: each element in the original matrix has been multiplied by 3. Hence, we arrive at the method for multiplying a matrix by a constant: you multiply each element by that constant. The resulting matrix has the same dimensions as the original.

Row Operations

Row operations are ways to change matrices. There are three types of row operations: row switching, that is interchanging two rows of a matrix; row multiplication, or multiplying all entries of a row by a non-zero constant; and finally row addition, which means adding a multiple of a row to another row. These row operations are used in a number of ways, including solving linear equations and finding inverses.

Source: https://www.boundless.com/algebra/systems-of-equations-and-matrices/matrix-operations/addition-and-subtraction-scalar-multiplication/

Matrix Multiplication

When multiplying matrices, the elements of the rows in the first matrix are multiplied with corresponding columns in the second matrix.

KEY POINTS

- If A is an n×m matrix and B is an m×p matrix, the result AB of their multiplication is an n×p matrix defined only if the number of columns m in A is equal to the number of rows m in B.
- Treating the rows and columns in each matrix as row and column vectors respectively, this entry is also their vector dot product.
- The product of a square matrix multiplied by a column matrix arises naturally in linear algebra for solving linear equations and representing linear transformations.

Assume two matrices are to be multiplied, the generalization to any number is discussed below. If A is an n×m matrix and B is an m×p **matrix**, the result AB of their multiplication is an n×p matrix defined only if the number of columns m in A is equal to the number of rows m in B.

General Definition

The arithmetic process of multiplying numbers in row i in matrix A and column j in matrix B, then adding to obtain entry ij in the final matrix.

When multiplying matrices, the elements of the rows in the first matrix are multiplied with corresponding columns in the second matrix. One may compute each entry in the third matrix one at a time. For two matrices,

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1m} \\ A_{21} & A_{22} & \dots & A_{2m} \end{pmatrix} \quad B = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1p} \\ B_{21} & B_{22} & \dots & B_{2p} \end{pmatrix}$$

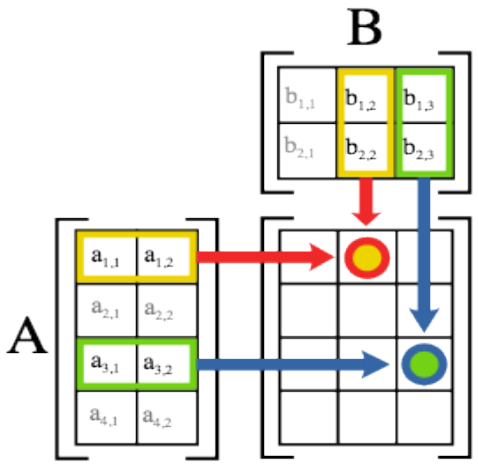
where necessarily the number of columns in A equals the number of rows in B equals m, the matrix product AB is defined (with no multiplication signs or dots), and where AB has entries defined by the equation:

$$A = \begin{pmatrix} (AB)_{11} & (AB)_{12} & \dots & (AB)_{1mp} \\ (AB)_{21} & (AB)_{22} & \dots & (AB)_{2mp} \end{pmatrix}$$

Treating the rows and columns in each matrix as row and column vectors respectively, this entry is also their vector dot product:

Usually the entries are numbers or expressions, but can even be matrices themselves, see block matrix. The matrix product can still be calculated exactly the same way.

Figure 6.11 Matrix Multiplication



This figure illustrates diagrammatically the product of two matrices A and B, showing how each intersection in the product matrix corresponds to a row of A and a column of B.

Illustration

Figure 6.11 illustrates diagrammatically the product of two matrices A and B, showing how each intersection in the product matrix corresponds to a row of A and a column of B.

$$\begin{bmatrix} a_{11} & a_{12} \\ \cdot & \cdot \\ a_{31} & a_{32} \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} \cdot & b_{12} & b_{13} \\ \cdot & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} \cdot & x_{12} & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & x_{33} \\ \cdot & \cdot & \cdot \end{bmatrix}$$

The values at the intersections marked with circles are:

$$x_{12} = (a_{11}, a_{12}) \cdot (b_{12}, b_{22}) = (a_{11}b_{12}) + (a_{12}b_{22})$$

$$x_{33} = (a_{31}, a_{32}) \cdot (b_{13}, b_{23}) = (a_{31}b_{13}) + (a_{32}b_{23})$$

Source: https://www.boundless.com/algebra/systems-of-equations-and-matrices/matrix-operations/matrix-multiplication/

Matrix Equations

Matrices can be used to compactly write and work with systems of multiple linear equations.

KEY POINTS

- If A is an m-by-n matrix, x designates a column vector (i.e., n×1-matrix) of n variables $x_1, x_2, \ldots, x_n$, and b is an m×1-column vector, then the matrix equation: $Ax = b$

Matrices can be used to compactly write and work with systems of equations.

A matrix is made of m rows and n columns. It is very common for the dimensions of a matrix to be written like m by n, or mxn.

This is an example of a 2x2 **matrix**.

$\begin{bmatrix} a \\ b \end{bmatrix}$ This is an example of a 2x1 matrix.

As we have learned in previous sections, matrices can be manipulated in any way that a normal equation can be manipulated in. This is very helpful when we start to work with systems of equations. It is helpful to understand how to use matrices to solve these systems before we start to just 'plug them into a calculator'.

Lets take the following system of equations:

3x+2y-z = 1

2x+2y+4z = -2

$x + \frac{1}{2}y - z = 0$

Solving a System of Equations Using a Matrix

It is possible to solve this system using the elimination or substitution method, but it would be easier to do it with a matrix operation. Before we start setting up the matrices, it is important to do the following:

- Make sure that all of the equations are written in a similar manner, meaning the variables need to all be in the same order
- Make sure that one side of the equation is only variables and their coefficients, and the other side is just constants.

There are 3 matrices that need to be set up. Matrix [A], which is made up of the variable coefficients. This matrix will have as many rows as variables, and as many columns as equations. Matrix [x], which is a matrix made up of just variables. This matrix will have as many rows as variables, but only one column. Matrix [b] is made up of the equation solutions, that is, the right side of the equations.

This matrix will have as many rows as variables, but only one column.

These matrices will be written as $Ax = b$. In order to solve the equations, use the following equation: $x = A^{-1} * b$. We will not cover how to get the inverse of a matrix in this section, because it is outside the scope of this atom. This can easily be plugged into a graphing calculator, or you can refer to that atom.

This is how we will write Ax=b for this system:

$$\begin{bmatrix} 3 & 2 & -1 \\ 2 & 2 & 4 \\ -1 & \frac{1}{2} & -1 \end{bmatrix} * \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$$

When you use the given equation, A-1(b)=x, you will be given a solution in this format:

$$\begin{bmatrix} 3 & 2 & -1 \\ 2 & 2 & 4 \\ -1 & \frac{1}{2} & -1 \end{bmatrix}^{-1} * \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -2 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Another example, if A is an m-by-n matrix, x designates a column vector (i.e., n×1-matrix) of n variables $x_1, x_2, \ldots, x_n$, and b is an m×1-column vector, then the matrix equation:

$$Ax = b$$

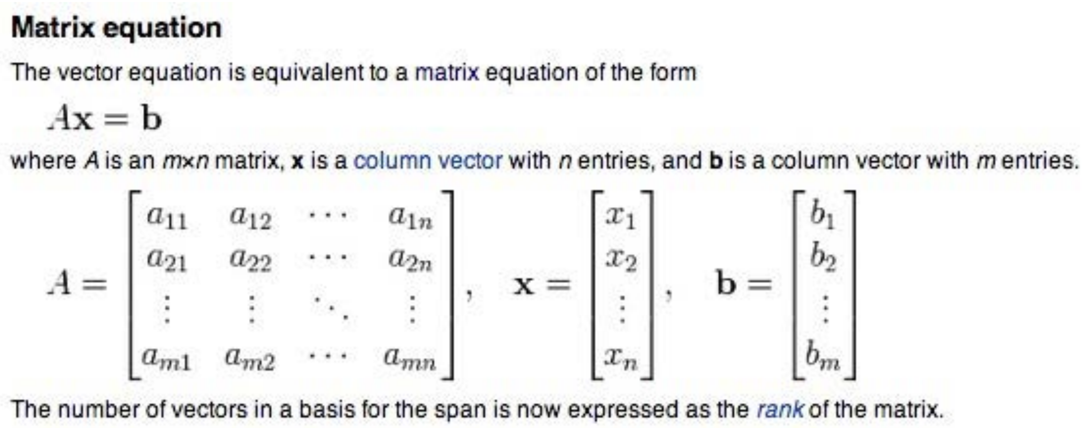
Figure 6.12 is equivalent to the following system of linear equations:

$$A_{1,1}x_1 + A_{1,2}x_2 + \ldots + A_{1,n}x_n = b_1$$

...

$$A_{m,1}x_1 + A_{m,2}x_2 + \ldots + A_{m,n}x_n = b_m$$

Figure 6.12 Matrix equation



Matrix equation

The vector equation is equivalent to a matrix equation of the form

$$\mathbf{Ax} = \mathbf{b}$$

where A is an $m \times n$ matrix, **x** is a column vector with n entries, and **b** is a column vector with m entries.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

The number of vectors in a basis for the span is now expressed as the *rank* of the matrix.

The vector equation is equivalent to a matrix equation of the form: Ax=b, where A is an m×n matrix, x is a column vector with n entries, and b is a column vector with m entries.

Section 5

Inverses of Matrices

https://www.boundless.com/algebra/systems-of-equations-and-matrices/inverses-of-matrices/

The Identity Matrix

The identity matrix [I] is defined so that [A][I]=[I][A]=[A], i.e. it is the matrix version of multiplying a number by one.

KEY POINTS

- For any square matrix, its identity matrix is a diagonal stretch of 1s going from the upper-left-hand corner to the lower-right, with all other elements being 0.
- Non-square matrices do not have an identity. That is, for a non-square matrix [A], there is no matrix such that [A][I]=[I][A]=[A].
- Proving that the identity matrix functions as desired requires the use of matrix multiplication.

When multiplying numbers, the number 1 has a special property: when multiplying 1 by any number, the same number back is gotten back. This idea can be expressed with the following property as an algebraic generalization: $1x = x$

The matrix that has this property is referred to as the **identity matrix**.

Definition of the Identity Matrix

The identity matrix, designated as $[I]$, is defined by the property: $[A][I] = [I][A] = [A]$.

Note that the definition of $[I]$ stipulates that the multiplication must commute—that is, it must yield the same answer no matter in which order multiplication is done. This stipulation is important because, for most matrices, multiplication does not commute.

What matrix has this property? The first guess might be a matrix full of 1s, but that does not work: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 7 \\ 3 & 7 \end{pmatrix}$ so

$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ is not an identity matrix.

The matrix that does work is a diagonal stretch of 1s, with all other elements being 0.

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

so $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity matrix for 2x2 matrices.

For a 3x3 matrix, the identity matrix is a 3x3 matrix with diagonal 1's and the rest 0's, for example

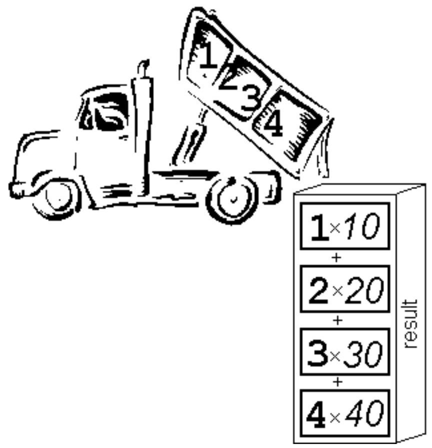
$$\begin{pmatrix} 2 & \pi & -3 \\ 5 & -2 & \frac{1}{2} \\ 9 & 8 & 8.3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & \pi & -3 \\ 5 & -2 & \frac{1}{2} \\ 9 & 8 & 8.3 \end{pmatrix}$$

so $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is the identity matrix for 3x3 matrices.

It is important to confirm those multiplications, and also confirm that they work in reverse order (as the definition requires). Remember, when doing matrix multiplication, multiply the first element of the first row of the first matrix by the first element of the first column of the second matrix, add it to the second by the second, etc., to determine the first element of the first row and column of the product matrix. The second element of the first row of the product matrix is the same thing, but done by multiplying the first row of the first matrix by the second column of the second matrix, then summing. In this way, each element of the product matrix can be determined. One mnemonic for remembering how to multiply matrices is that it is similar to taking a row of the first matrix, putting it into a dump truck, and tilting it into a column of the second matrix, then multiplying whatever is next to each other and adding the products ([*Figure 6.13*](#)). The final location of the element will be where the row and column intersect.

Why is there no identity for a non-square matrix? There is no identity for a non-square matrix, because of the requirement of commutativity. For a non-square matrix [A] one might be able to find a matrix $[I]$ such that $[A][I] = [A]$. However, if the order is reversed, then an illegal multiplication will be left. The reason for this is because, for two matrices to be multiplied together, the first matrix must have the same number of columns as the second has rows.

Figure 6.13 Matrix multiplication



One way to remember how to multiply matrices is to imagine a dump truck taking the row of the first matrix and "dumping" it next to the column of the second. Multiply each pair and add the products together to get the element of the product matrix.

The Inverse of a Matrix

The inverse of matrix [A] is $[A]^{-1}$, and is defined by the property: $[A][A]^{-1}=[A]^{-1}[A]=[I]$.

KEY POINTS

- Note that, just as in the definition of the identity matrix, this definition requires commutativity—the multiplication must work the same in either order.
- To be invertible, a matrix must be square, because the identity matrix must be square as well.
- To determine the inverse of the matrix $\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}$, set $\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Then solve for a, b, c, and d.

Having seen that the number 1 plays a special role in multiplication, because $1x = x$, the inverse of a number is defined as the number that multiplies by that number to give 1. b is the inverse of a if the inverse of a matrix multiplies by that matrix to give the identity matrix.

Definition of Inverse Matrix

The inverse of matrix $[A]$, designated as $[A]^{-1}$, is defined by the property: $[A][A]^{-1} = [A]^{-1}[A] = [I]$, where $[I]$ is the identity matrix.

Note that, just as in the definition of the identity matrix, this definition requires commutativity—the multiplication must work the same in either order.

Note also that only square matrices can have an inverse. The definition of an **inverse matrix** is based on the identity matrix $[I]$, and we already said that only square matrices even have an identity.

The method for finding an inverse matrix comes directly from the definition, along with a little algebra.

Finding an Inverse Matrix

Find the inverse of $\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}$.

Set $\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

This is the key step. It establishes $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ as the inverse that is being looked for, by asserting that it fills the definition of an inverse matrix: when multiplying this mystery matrix by our original matrix, the result is I. When solving for the four variables a, b, c, and d, then the inverse of the matrix will be found.

Next, do the multiplication giving $\begin{pmatrix} 3a+4c & 3b+4d \\ 5a+6c & 5b+6d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Make sure to check this step, it provides great practice. For instance, start by multiplying the first row times the first column, and the result is $3a+4c$.

Therefore, one can set:

$3a+4c=1$

$3b+4d=0$

$5a+6c=0$

$5b+6d=1$

For two matrices to be equal: every element in the left must equal its corresponding element on the right. So, for these two matrices to equal each other, all four of these equations must hold.

Therefore, the results are:

$a=-3$

$b=2$

$c=2\frac{1}{2}$

$d=-1\frac{1}{2}$

Solve the first two equations for a and c by using either elimination or substitution. Solve the second two equations for b and d by using either elimination or substitution.

So, the inverse is $\begin{pmatrix} -3 & 2 \\ 2\frac{1}{2} & -1\frac{1}{2} \end{pmatrix}$

Solving for the four variables results in the inverse.

If an inverse has been found, then when it is multiplied by the original matrix, then the result should be:

$$\begin{pmatrix} -3 & 2 \\ 2\frac{1}{2} & -1\frac{1}{2} \end{pmatrix}\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

After multiplying, one gets

$$\begin{pmatrix} (-3)(3)+(2)(5) & (-3)(4)+(2)(6) \\ (2\frac{1}{2})(3)_{+}(-1\frac{1}{2})(5) & (2\frac{1}{2})(4)+(-1\frac{1}{2})(6) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

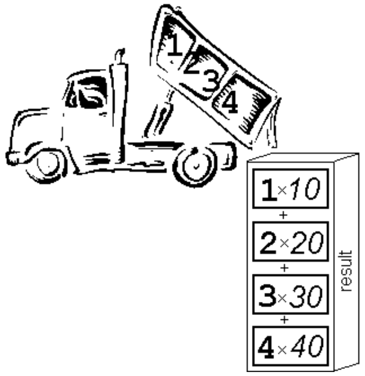
Simplifying the problem gives $\begin{pmatrix} -9+10 & -12+12 \\ 7\frac{1}{2}-7\frac{1}{2} & 10-9 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Therefore, the inverse works.

Note that, to fully test it, one has to try the multiplication in both orders, because, in general, changing the order of a matrix multiplication changes the answer. The definition of an inverse matrix specifies that it must work both ways. Only one order was shown above, so technically, this inverse has only been half-tested.

This process does not have to be memorized: it should make logical sense. Everything one knows about matrices should make logical sense, except for the very arbitrary-looking definition of matrix multiplication. If it helps, remember the dump truck mnemonic (*Figure 6.14*).

Figure 6.14 Multiplying matrices



Remember, when multiplying matrices, that one mnemonic to remember which terms to multiply together is to pretend that you have placed the row from the first matrix onto a dump truck, and are "dumping" the terms onto the column from the second matrix.

Solving Systems of Equations Using Matrices

A system of equations can be readily solved using the concepts of the inverse matrix and matrix multiplication.

KEY POINTS

- Using matrices to solve systems of equations can drastically reduce the workload on you. Consider the following three equations:
  $x + 2y - z = 11$
  $2x - y + 3z = 7$
  $7x - 3y - 2z = 2$
- To solve these equations using matrices, we first define a 3×3 matrix [A], which is the coefficients of all the variables on the left side of the equal signs: $[A] = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ 7 & -3 & -2 \end{bmatrix}$ Also, define a 3×1 matrix [B], which is the numbers on the right side of the equal signs: $[B] = \begin{bmatrix} 11 \\ 7 \\ 2 \end{bmatrix}$.
- In order to determine the values of x, y, and z, we simply multiply the inverse of [A] times [B]. This is most readily done using a calculator. The calculator responds with a 3×1 matrix, which is all three answers. In this case, x=3, y=5, and z=2.

At this point, you may be left with a pretty negative feeling about matrices. The initial few ideas—adding matrices, subtracting them, multiplying a matrix by a constant, and matrix equality—seem almost too obvious to be worth talking about. On the other hand, multiplying matrices and taking determinants seem to be strange, arbitrary sequences of steps with little or no purpose.

A great deal of it comes together in solving **linear equation**s. We have seen, in the chapter on simultaneous equations, how to solve two equations with two unknowns. But suppose we have three equations with three unknowns? Or four, or five? Such situations are more common than you might suppose in the real world. And even if you are allowed to use a calculator, it is not at all obvious how to solve such a problem in a reasonable amount of time.

Surprisingly, the things we have learned about matrix multiplication, about the identity matrix, about inverse matrices, and about matrix equality, all give us a very fast way to solve such problems on a calculator!

Consider the following example, three equations with three unknowns:

$x + 2y - z = 11$

$2x - y + 3z = 7$

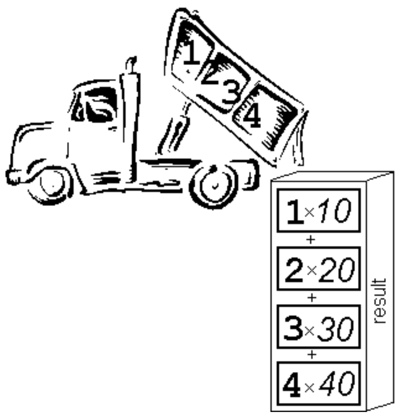
$7x - 3y - 2z = 2$

Define a 3×3 matrix [A], which is the coefficients of all the variables on the left side of the equal signs: $[A] = \begin{bmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ 7 & -3 & -2 \end{bmatrix}$

Define a 3×1 matrix [B], which is the numbers on the right side of the equal signs: $[B] = \begin{bmatrix} 11 \\ 7 \\ 2 \end{bmatrix}$

Now, in order to determine the values of x, y, and z, we simply multiply the inverse of [A] times [B]. This can be done by hand, finding the **inverse matrix** of [A], then performing the appropriate matrix multiplication (*Figure 6.13*). However, if you have a graphing calculator, the situation is much easier. Punch these matrices into your calculator, and then ask the calculator for $[A^{-1}][B]$, that is, the inverse of matrix [A], multiplied by matrix [B].

Figure 6.15 Matrix multiplication



When doing matrix multiplication, it may help to remember the mnemonic device of taking the rows of the first matrix and "dumping" them into the columns of the second matrix.

The calculator responds with a

3×1 matrix that is all three answers. In this case, x=3, y=5, and z=2.

The whole process takes no longer than it does to punch a few matrices into the calculator. And it works just as quickly for 4 equations with 4 unknowns, or 5, etc.

Huh? Why the heck did that work?

Solving linear equations in this way is fast and easy. The proof, unfortunately, is beyond the scope of this atom.

Section 6

Determinants and Cramer's Rule

Determinants of Square Matrices

The determinant of a square matrix is computed by recursively performing the Laplace expansion to find the determinant of smaller matrices.

KEY POINTS

- The determinant of a 2-by-2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is defined to be $ad - bc$.
- The Laplace expansion for a n-by-n square matrix B is $\sum_{j'=1}^{n} b_{ij'}C_{ij'}$ for some fixed i or $\sum_{i'}^{n} b_{i'j}C_{i'j}$ for some fixed j.
- Recursively performing the Laplace expansion on each of the smaller sub-matrices in the Cofactor will eventually produce a small enough sub-matrix for which the determinant is known.

In linear algebra, the **determinant** is a value associated with a square matrix. It can be computed from the entries of the matrix by a specific arithmetic expression, while other ways to determine its value exist as well. The determinant provides important information when the matrix is, that of the coefficients of a system of linear equations, or when it corresponds to a linear transformation of a vector space: in the first case the system has a unique solution if and only if the determinant is nonzero, while in the second case that same condition means that the transformation has an inverse operation. A geometric interpretation can be given to the value of the determinant of a square matrix with real entries: the absolute value of the determinant gives the scale factor by which area or volume is multiplied under the associated linear transformation, while its sign indicates whether the transformation preserves orientation. Thus a 2×2 matrix with determinant -2, when applied to a region of the plane with finite area, will transform that region into one with twice the area, while reversing its orientation.

Determinants occur throughout mathematics. The use of determinants in calculus includes the Jacobian determinant in the substitution rule for integrals of functions of several variables. They are used to define the characteristic polynomial of a matrix that is an essential tool in eigenvalue problems in linear algebra. In some cases, they are used just as a compact notation for expressions that would otherwise be unwieldy to write down.

The determinant of a matrix A is denoted det(A), det A, or |A|. In the case where the matrix entries are written out in full, the

determinant is denoted by surrounding the matrix entries by vertical bars instead of the brackets or parentheses of the matrix.

For instance, the determinant of the matrix $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is written

$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ and has the value $aei + bfg + cdh - afh - bdi - ceg$.

Determinant of a 2-by-2 Matrix

For a 2-by-2 matrix, $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ is defined to be ad-bc.

If the matrix entries are real numbers, the matrix A can be used to represent two linear mappings: one that maps the standard basis vectors to the rows of A, and one that maps them to the columns of A. In either case, the images of the basis vectors form a parallelogram that represents the image of the unit square under the mapping. The parallelogram defined by the rows of the above matrix is the one with vertices at (0,0), (a,b), (a + c, b + d), and (c,d), as shown in *Figure 6.16*. The absolute value of is the area of the parallelogram, and thus represents the scale factor by which areas are transformed by A.

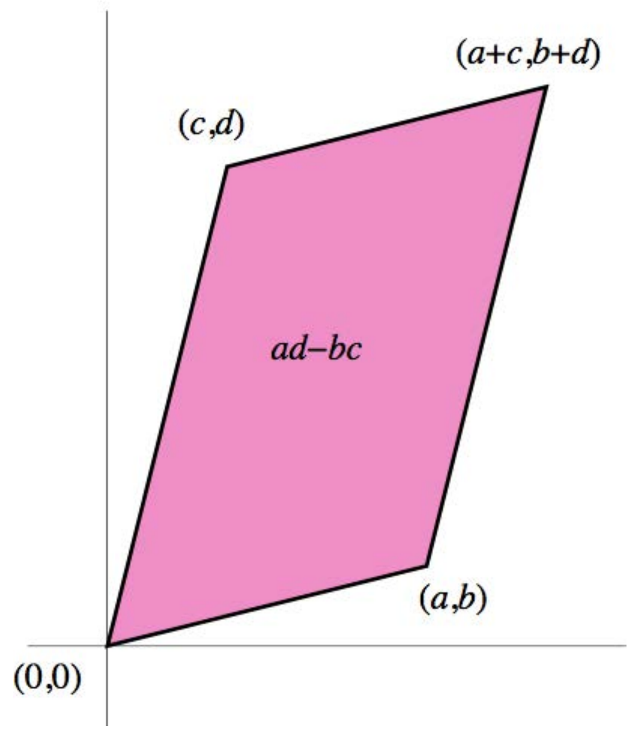


Figure 6.16 Determinant as Area

The area of the parallelogram is the absolute value of the determinant of the matrix formed by the vectors representing the parallelogram's sides.

Laplace Expansion

In linear algebra, the Laplace expansion, named after Pierre-Simon Laplace, also called **cofactor** expansion, is an expression for the determinant |B| of an n × n square matrix B that is a weighted sum of the determinants of n sub-matrices of B, each of size (n−1) × (n −1). The Laplace expansion is of theoretical interest as one of several ways to view the determinant, as well as of practical use in determinant computation.

The i, j cofactor of B is the scalar C_{ij} defined by: $C_{ij} = (-1)^{i+j} M_{ij}$ where M_{ij} is the i, j minor matrix of B, that is, the determinant of the

(n-1)x(n-1) matrix that results from deleting the i-th row and j-th column of B.

Then the Laplace expansion is given by the following:

Suppose $B = (b_{ij})$ is an n x n matrix and keep one of i, j $\in$ {1, 2,..., n} fixed.

Then its determinant |B| is given by:

$$|B| = b_{i1}C_{i1} + b_{i2}C_{i2} + \ldots + b_{in}C_{in} = \sum_{j'=1}^{n} b_{ij'}C_{ij'}$$

Or:

$$|B| = b_{1j}C_{1j} + b_{2j}C_{2j} + \ldots + B_{nj}C_{nj} = \sum_{i'=1}^{n} b_{i'j}C_{i'j}$$

Determinant of a 3-by-3 Matrix

Using the Laplace expansion to find the determinant of a 3-by-3 matrix when given a general 3-by-3 matrix $B = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$, choose to fix a row i or a column j. Choose the first row, and fix i=1.Thus the determinant is:

$$|B| = b_{11}C_{11} + b_{12}C_{12} + b_{13}C_{13}$$

$$= a(-1)^{1+1}M_{11} + b(-1)^{1+2}M_{12} + c(-1)^{1+3}M_{13}$$

$$= aM_{11} - bM_{12} + cM_{13}$$

$$= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= a(ei - hf) - b(di - fg) + c(dh - eg)$$

$$= aei + bfg + cdh - ahf - bdi - ceg$$

Laplace Expansion as Recursive Function

The Laplace expansion can be used as a **recursive** function for any arbitrary n-by-n matrix. Start with n x n matrix, and fix either a row or column and perform the Laplace expansion to get n sub-matrices of size (n-1) x (n-1). Then recurse on each of those (n-1) x (n-1) sub-matrices to get (n-1) sub-matrices of size (n-2) x (n-2), for EACH (n-1) sub-matrix. Now there are n*(n-1) sub-matrices of size (n-2) x (n-2). Then perform the Laplace expansion on each of those sub-matrices, until eventually ending up with a sub-matrix of a size that the determinant is known, for instance a 3-by-3 matrix. However, to find the determinant of a very large matrix becomes rather difficult quickly. Indeed there will end up being n! sub-matrices of size 1, and the determinant of a matrix of size 1 is just the single element of that matrix.

Source: https://www.boundless.com/algebra/systems-of-equations-and-matrices/determinants-and-cramer-s-rule/determinants-of-square-matricies/

Cofactors

The cofactor of an entry (i, j) of a matrix A is the signed minor of that matrix.

KEY POINTS

- Let A be an m × n matrix and k an integer with 0 < k ≤ m, and k ≤ n. A k × k minor of A is the determinant of a k × k matrix obtained from A by deleting m − k rows and n − k columns.
- The first minor of a matrix M_{ij} is formed by removing the ith row and jth column of the matrix, and retrieving the determinant of the smaller matrix.
- The cofactor of an element a_{ij} of a matrix A, written as C_{ij} is defined as $(-1)^{i+j} M_{ij}$.

In linear algebra, the cofactor (sometimes called adjunct) describes a particular construction that is useful for calculating both the determinant and inverse of square matrices. Specifically the cofactor of the (i, j) entry of a matrix, also known as the (i, j) cofactor of that matrix, is the signed minor of that entry.

Minor

To know what the signed **minor** is, we need to know what the minor of a matrix is. In linear algebra, a minor of a matrix A is the

determinant of some smaller square matrix, cut down from A by removing one or more of its rows or columns. Minors obtained by removing just one row and one column from square matrices (first minors) are required for calculating matrix cofactors, which in turn are useful for computing both the determinant and inverse of square matrices.

Definition

Let A be an m × n matrix and k an integer with 0 < k ≤ m, and k ≤ n. A k × k minor of A is the determinant of a k × k matrix obtained from A by deleting m − k rows and n − k columns.

More often than not you'll only remove one column and row at a time, and so the (i, j) minor, often denoted M_{ij}, of an $n \times n$ square matrix A is defined as the determinant of the $(n-1) \times (n-1)$ matrix formed by removing from A its ith row and jth column. An (i, j) minor is also referred to as the (i, j)th minor, or simply i, j minor.

M_{ij} is also called the minor of the element a_{ij} of matrix A.

A minor that is formed by removing only one row and column from a square matrix A (such as M_{ij}) is called a first minor. When two rows and columns are removed, this is called a second minor.

Example

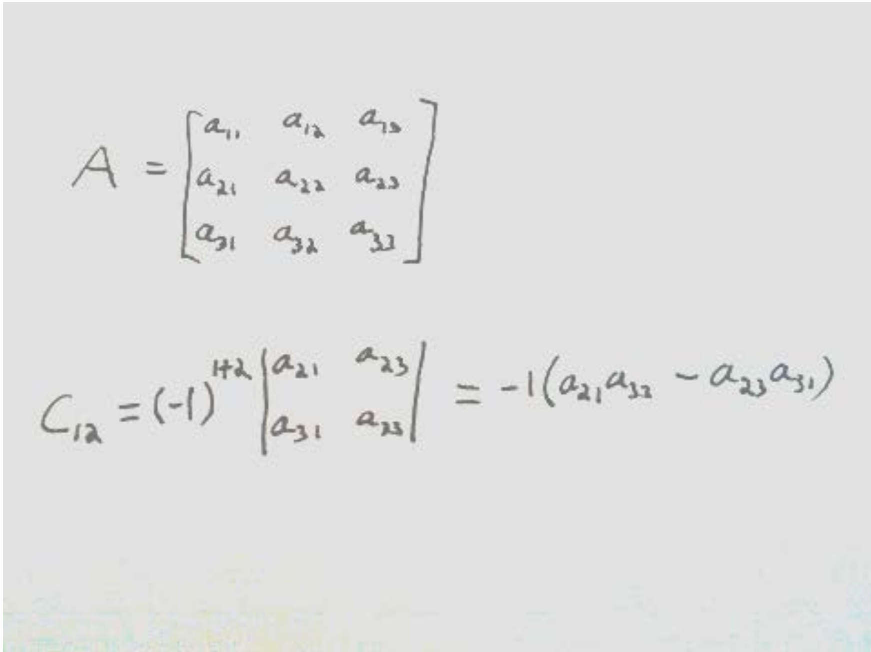
For example, say we are given the matrix:

$$\begin{bmatrix} 1 & 4 & 7 \\ 3 & 0 & 5 \\ -1 & 9 & 11 \end{bmatrix}$$

The minor M_{23} is the determinant of the 2 x 2 matrix formed by removing the 2nd row and 3rd column, i.e.

$$\begin{vmatrix} 1 & 4 & \bullet \\ \bullet & \bullet & \bullet \\ -1 & 9 & \bullet \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ -1 & 9 \end{vmatrix} = (9 - (-4)) = 13$$

Figure 6.17 Cofactor



Here is a cofactor of an arbitrary 3x3 matrix taking the first row and second column out.

Where a black dot represents an element we are removing.

Cofactor

The cofactor of a_{ij} entry of a matrix is defined as:

$C_{ij} = (-1)^{i+j} M_{ij}$

Informal Approach

Finding the minors of a matrix A is a multi-step process:

1. Choose an entry a_{ij} from the matrix.
2. Cross out the entries that lie in the corresponding row i and column j.
3. Rewrite the matrix without the marked entries.
4. Obtain the determinant of this new matrix.

M_{ij} is termed the minor for entry a_{ij}.

If i + j is an even number, the cofactor of coincides with its minor: $C_{ij} = M_{ij}$

Otherwise, it is equal to the additive inverse of its minor: $C_{ij} = - M_{ij}$

Cramer's Rule

Cramer's Rule uses determinants to solve for a solution to the equation Ax=b, when A is a square matrix.

KEY POINTS

- Cramer's Rule only works on square matrices that have a non-zero determinant and a unique solution.
- Cramer's Rule is defined as $x_i = \frac{det(A_i)}{det(A)} \quad i = 1,\ldots,n$, where A_i is the matrix formed by replacing the ith column of A by the column vector b in the equation.
- Cramer's Rule is efficient for solving small systems and can be calculated quite quickly; however, as the system grows, calculating the new determinants can be tedious.

Determinants

Determinants play a critical role in the application of Cramer's Rule. A determinant is a value associated with a **square matrix**. It can be computed from the entries of the matrix by a specific arithmetic expression, while other ways to determine its value exist as well. The determinant provides important information when the matrix is that of the coefficients of a system of linear equations, or when it corresponds to a linear transformation of a vector space: in the first case the system has a unique solution if and only if the determinant is nonzero, while in the second case that same condition means that the transformation has an inverse operation.

A geometric interpretation can be given to the value of the determinant of a square matrix with real entries: the absolute value of the determinant gives the scale factor by which area or volume is multiplied under the associated linear transformation, while its sign indicates whether the transformation preserves orientation. Thus a 2×2 matrix with determinant -2, when applied to a region of the plane with finite area, will transform that region into one with twice the area, while reversing its orientation.

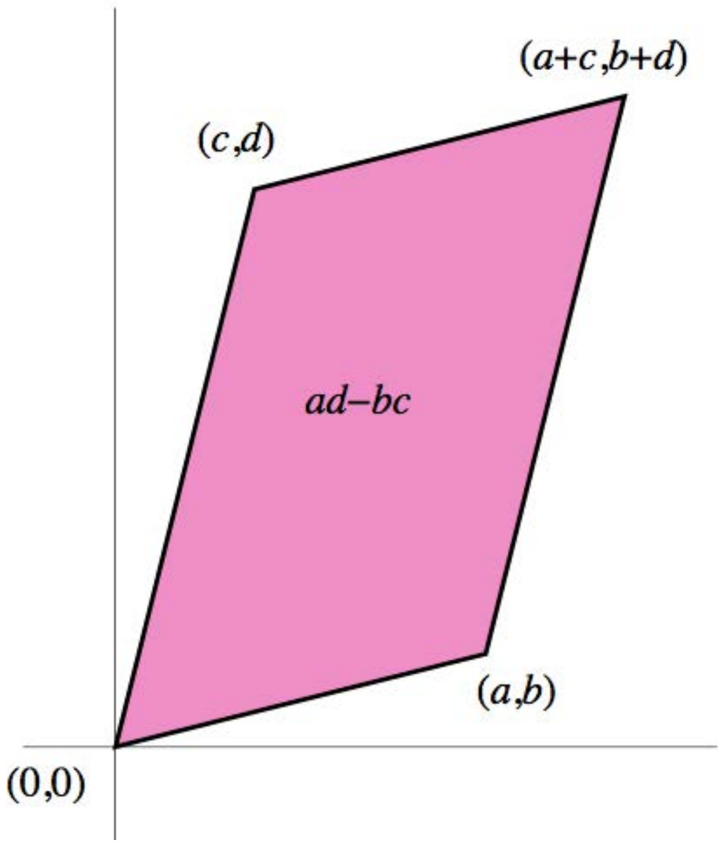
The determinant of a matrix A is denoted det(A), det A, or |A|. In the case where the matrix entries are written out in full, the determinant is denoted by surrounding the matrix entries by vertical bars instead of the brackets or parentheses of the matrix. For instance, the determinant of the matrix

$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is written $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$ and has the value :

$$aei + bfg + cdh - bdi - afh.$$

A simple 2x2 matrix: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

Figure 6.18 Determinant as Area



The area of the parallelogram is the absolute value of the determinant of the matrix formed by the vectors representing the parallelogram's sides.

The images of the basis vectors form a parallelogram that represents the image of the unit square under the mapping (*Figure 6.18*). The parallelogram defined by the rows of the above matrix is the one with vertices at (0,0), (a,b), (a + c, b + d), and (c,d), as shown in the accompanying diagram. The absolute value of is the area of the parallelogram, and thus represents the scale factor by which areas are transformed by A. The parallelogram formed by the columns of A is in general a different parallelogram. However, since the determinant is symmetric with respect to rows and columns, the area will be the same.

Cramer's Rule

Cramer's Rule is an explicit formula for the solution of a system of linear equations with as many equations as unknowns, i.e. a square matrix, valid whenever the system has a unique solution. It expresses the solution in terms of the determinants of the (square) coefficient matrix and of matrices obtained from it by replacing one column by the vector of right hand sides of the equations.

Consider a system of n linear equations for n unknowns, represented in matrix multiplication form as follows:

$$Ax = b$$

where the n by n matrix A has a nonzero determinant, and the vector $Ex = (x_1, \ldots, x_n)^T$ is the column vector of the variables.

Then the theorem states that in this case the system has a unique solution, whose individual values for the unknowns are given by:

$$x_i = \frac{det(A_i)}{det(A)} \quad i = 1,\ldots,n$$

Where A_i is the matrix formed by replacing the ith column of Aby the column vector b.

Examples

Consider the linear system $\begin{cases} ax+by &= e \\ cx+dy &= f \end{cases}$, which in matrix format is $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix}$.

Assume the determinant is non-zero. Then, x and y can be found by Cramer's rule as

$$x = \begin{vmatrix} e & b \\ f & d \end{vmatrix} / \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \frac{ed - bf}{ad - bc}$$

and

$$y = \begin{vmatrix} a & e \\ c & f \end{vmatrix} / \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \frac{af - ec}{ad - bc}$$

The rules for a 3x3 matrix are similar. Given $\begin{cases} ax+by+cz &= j \\ dx+ey+fz &= k, \\ gx+hy+iz &= l \end{cases}$

which in matrix format is $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} j \\ k \\ l \end{bmatrix}$.

Then the values of x, y and z can be found as follows:

$$x = \frac{\begin{vmatrix} j & b & c \\ k & e & f \\ l & h & i \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}} y = \frac{\begin{vmatrix} a & j & c \\ d & k & f \\ g & l & i \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}} and \quad z = \frac{\begin{vmatrix} a & b & j \\ d & e & k \\ g & h & l \end{vmatrix}}{\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}}$$

Section 7

Systems of Inequalities and Linear Programming

https://www.boundless.com/algebra/systems-of-equations-and-matrices/systems-of-inequalities-and-linear-programming/

Graphs of Linear Inequalities

Graphing linear inequalities involves graphing the original line, and then shading in the area connected to the inequality.

KEY POINTS

- To graph a single linear inequality, first graph the inequality as if it were an equation. If the sign is ≤ or ≥, graph a normal line. If it is > or <, then use a dotted or dashed line. Then, shade either above or below the line, depending on if y is greater or less than mx + b.
- If there are multiple linear inequalities, then where all the shaded areas of each inequality overlap is the solutions to the system.
- If the shaded areas of all inequalities in a system do not overlap, then the system has no solution.

Single Inequality in Two Variables

The simplest **inequality** to graph is a single inequality in two variables, usually of the form

$$y \leq mx + b$$

where the inequality can be of any type, less than, less than or equal to, greater than, greater than or equal to, or not equal to. Graphing an inequality is easy. First, graph the inequality as if it were an equation. If the sign is ≤ or ≥, graph a normal line. If it is > or <, then use a dotted or dashed line. Then, shade either above or below the line, depending on if y is greater or less than mx + b. As an example, say $y \leq x + 2$. Since the equation is less than or equal to, start off by drawing the line $y = x + 2$. Next, note that y is less than or equal to x+2, which means that y can take on the values along the line, or any values below the line, and so we shade in all the values under the line to get *Figure 6.19*.

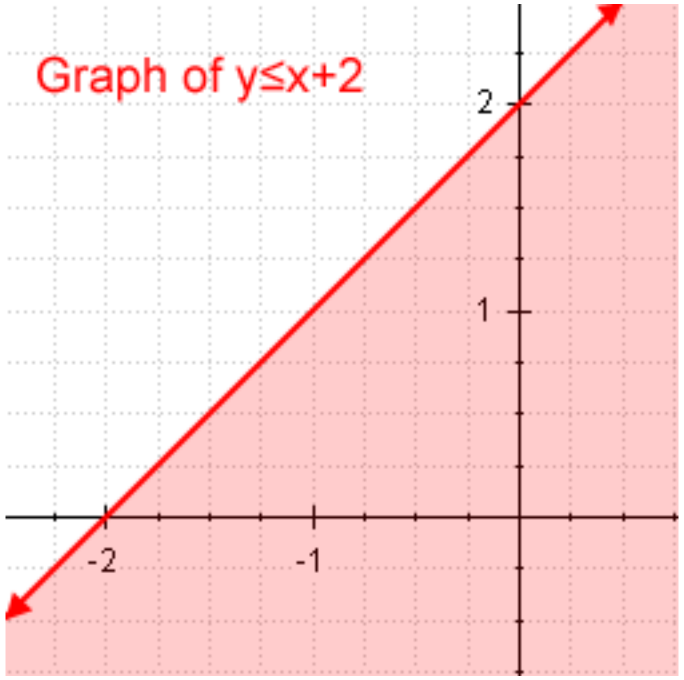


Figure 6.19 Graph of Single Inequality

Here is the graph of the single inequality y <= x + 2.

Multiple Inequalities in Two Variables

Now if there is more than one inequality, start off by graphing them one at a time, just as was done with a single inequality. To find

solutions for the group of inequalities, observe where the area of all of the inequalities overlap. These overlaps are solutions to the system. This also means that if there are inequalities that don't overlap, then there is no solution to the system.

To use a real world example, such as gas mileage:

The gas mileage sticker on cars gives two numbers: one for city driving, and one for highway driving. If a sticker says the car gets 25 mpg in the city and 32 mpg on the highway how far can you drive?

The abbreviation mpg stands for miles per gallon. The sticker on our car predicts that the car gets between 25 and 32 mpg, but when the car is driven one estimates how far the car is going and how much gas is in the tank. When graphing change the inequality to a function where x is the number of gallons of gas in our tank: $25x < f(x) < 32x$.

Figure 6.20 has two lines one for $y=25x$ and one for $y = 32x$. The yellow portion of the picture represents how far the car may be able to drive when there are x gallons of gas. The vertical line at 10 shows that it can drive between 250 and 320 miles on 10 gallons of gas. The difference between these two numbers is 70. If a line was drawn at 1 gallon of gas, it would be seen that the car could drive between 25 and 32 miles, and the difference would be 7.

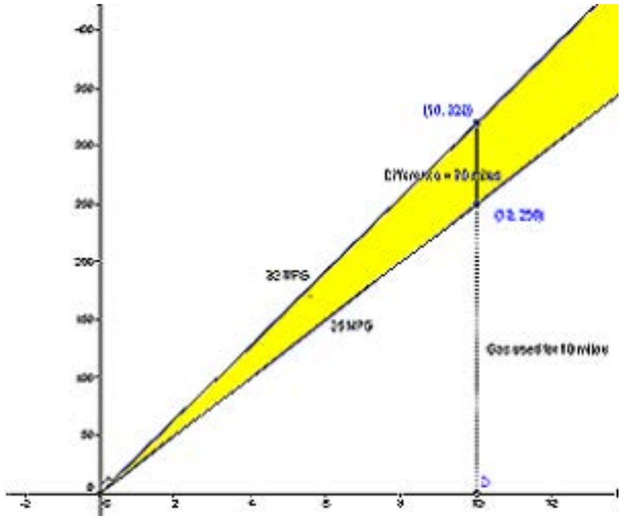


Figure 6.20 Gas Mileage Example

Here is an example showing in what range a car can drive given a city mpg and a highway mpg.

What this graph shows is the predicted range of miles a car can drive on a given amount of gas. The more gas a car has, the more likely the actual mileage is going to be different from what the sticker on the car predicted. If the actual mileage falls outside of this range than the car may need to go to the mechanic to make sure everything is running correctly. If the mileage is too high the odometer on the car may be broken. If the mileage is too low the car might need a tune up.

Solving Systems of Linear Inequalities

Solving for a system of linear inequalities requires finding values for each of the variables such that all equations are satisfied.

KEY POINTS

- To solve a system graphically, draw and shade in each of the inequalities on the graph, and then look for an area in which all of the inequalities overlap, this area is the solution.
- If there is no area in which all of the inequalities overlap, then the system has no solution.
- To solve a system non-graphically, find the intersection points, and then find out relative to those points which values still hold for the inequality. Narrow down these values until mutually exclusive ranges (no solutions) are found, or not, in which the solution is within your final range.

Graphical Method

To solve a system of linear inequalities, if possible, the easiest way to do this is by graphing. However, graphing is only possible if there are two or three variables. For two variables, when graphing, first draw all of the lines of the inequalities as if they were an equation, drawing a dotted line if it is either $<$ or $>$, and a solid line if it is either $\leq$ or $\geq$. After the line has been drawn, shade in, or indicate with hash marks, the area that corresponds to the inequality. For instance, if it is $<$ or $\leq$, shade in the area below the line. If it is $>$ or $\geq$ shade in the area above the line.

Once all of the inequalities have been drawn and shaded in, to find solutions to the system one needs to find areas in which all of the inequalities overlap each other. For example, given the system:

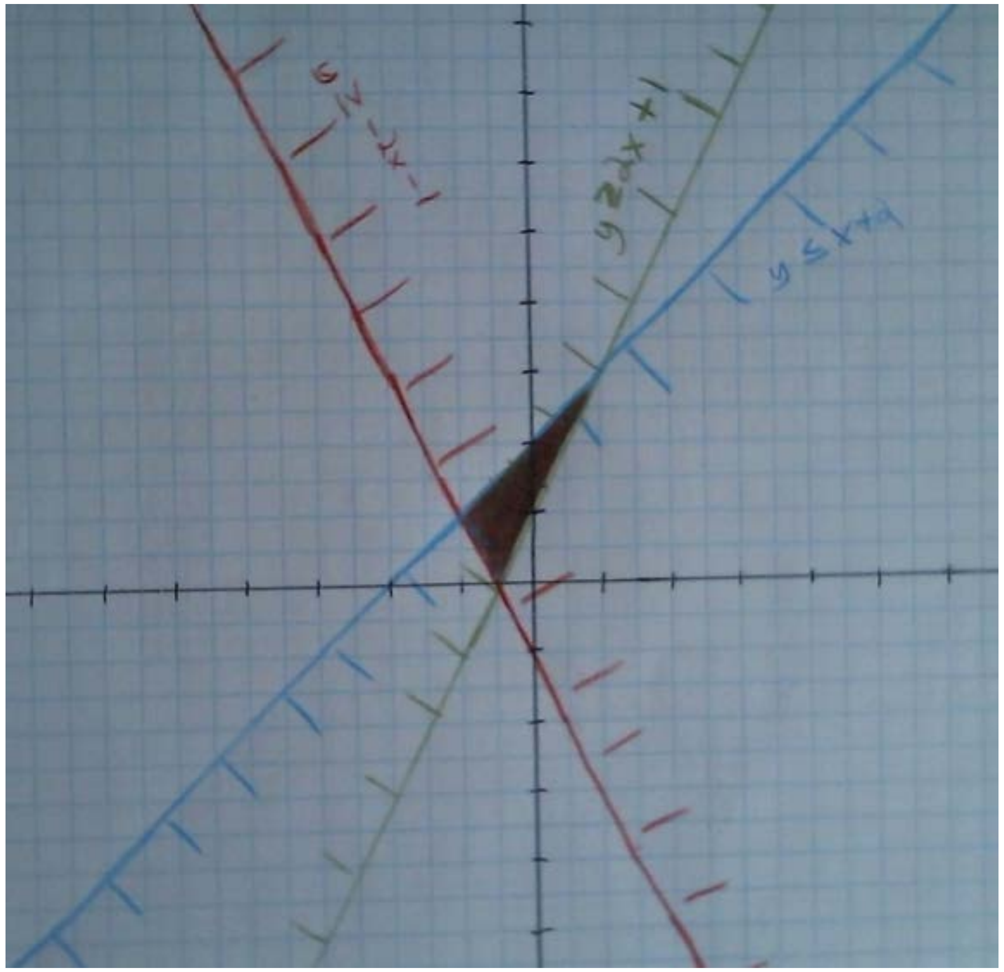
$$\begin{cases} y \geq & -2x - 1 \\ y \geq & 2x + 1 \\ y \leq & x + 2 \end{cases}$$

Draw each of the lines and shade in, or indicate, their corresponding inequalities, and then look to see what parts overlap. As can be seen in *Figure 6.21*, the shaded part in the middle is where all three inequalities overlap.

If all of the inequalities of a system fail to overlap over the same area, then there is no solution to that system. For instance, given the following system:

$$\begin{cases} y \geq & 1 \\ y \geq & 2x + 2 \\ y \leq & x + 1 \end{cases}$$

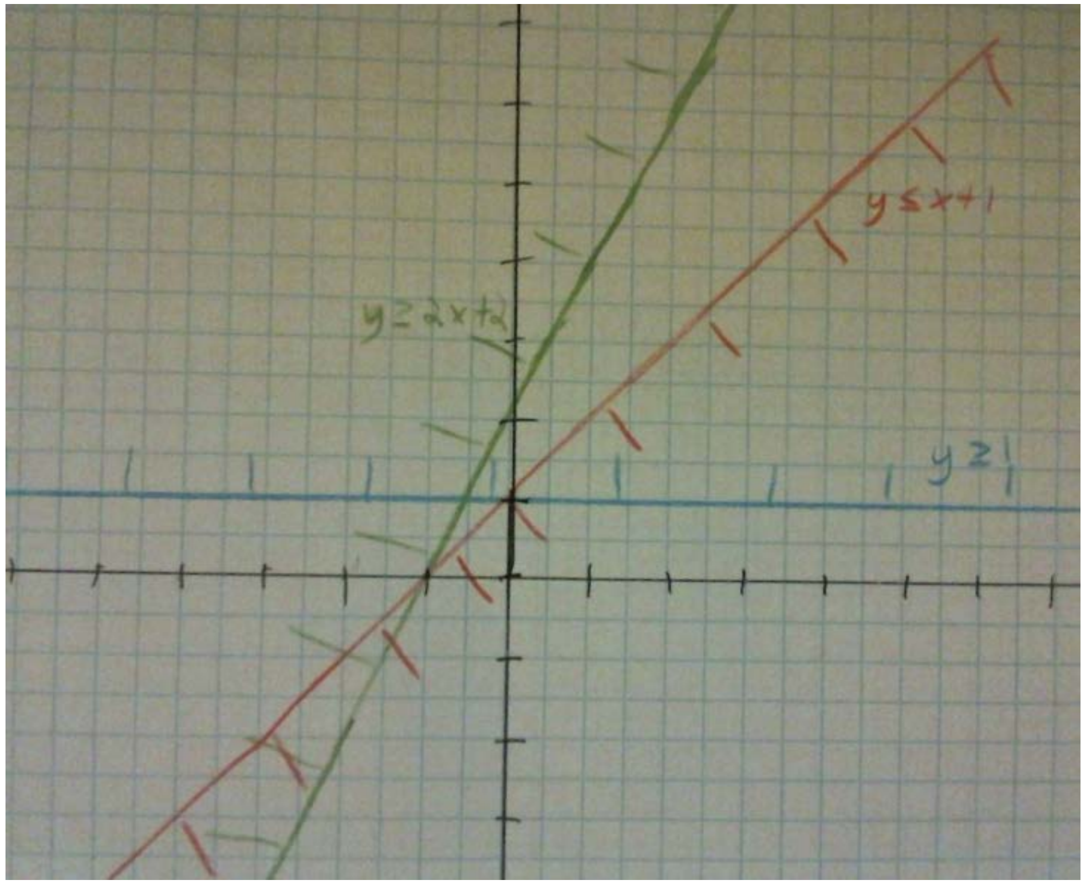
Figure 6.21 Solution to Linear Inequality System



This is a graph showing the solutions to a linear inequality system. Note that it is the overlapping areas of all three linear inequalities.

Again, draw all the inequalities and shade in, or indicate somehow, the area that the inequality covers. Notice that in this graph *Figure 6.22*, there is no part of the graph where all three inequalities overlap. There are plenty of areas where two of the three overlap at a time, but that is not enough, all three must overlap for those points to be a solution to the system.

Figure 6.22 Linear Inequality System With No Solutions



This is a graph showing a system of linear inequalities that has no solution as there is no point in which the areas of all three inequalities overlap.

Non-Graphical Method

Sometimes one may not wish to graph the equations, or simply cannot due to the number of variables. In this situation, find intervals in which certain variables satisfy the system by looking at two equations at a time. First, find the intersection point(s) of two of the equations, if there is no intersection, then the two inequalities are either **mutually exclusive**, or one of the inequalities is a

subset of the other. For a simple example, x > 2 and x < 1 are mutually exclusive, whereas x > 2 and x > 1 has x > 2 as a subset of x > 1. If they are mutually exclusive, then there is no solution.

Once an intersection point is found, determine on which side(s) of the intersection point the inequalities hold. For example, if there are two equations $y \geq -2x$ and $y \leq x+1$, the intersection is found to occur at $x = \frac{-1}{3}$.

Look at any point greater than that x value (say 0) to see that the first equation gives $y \geq 0$, and the second gives $y \leq 1$. Since these two equations are not mutually exclusive, these two equations are satisfied for any $x \geq \frac{-1}{3}$.

If the entire system had a third equation as well: $y \leq 4$, then next find the intersection points between this new equation and the other two. The intersection between $y \leq 4$ and $y \geq -2x$ occurs at $x = -2$, and that it is satisfied when $x \geq -2$. Note that the $x \geq \frac{-1}{3}$ found earlier is more restrictive.

Therefore, ignore the new inequality for x. Now the last set $y \leq 4$ and $y \leq x+1$ intersects at $x = 3$, further, by looking at x=2 and x=4, it is seen that for these two inequalities, $x \leq 3$. Now for the system to hold true, $\frac{-1}{3} \leq x \leq 3$, and to find valid y values, simply choose any x value in that range, and plug them in to get the range on the y values.

The non-graphical method is much more complicated, and is perhaps much harder to visualize all the possible solutions for a system of inequalities. However, having too many equations or too many variables, may be the only feasible method.

Application of Systems of Inequalities: Linear Programming

Linear programming involves finding an optimal solution for a linear equation given a number of constraints.

KEY POINTS

- The standard form for a linear program is: minimize $c \cdot x$, subject to $Ax = b, x_i \geq 0$. c is the coefficients of the objective function, x is the variables, A is the left-side of the constraints and b is the right side.
- The Simplex Method involves choosing an entering variable from the nonbasic variables in the objective function, finding the corresponding leaving variable that maintains feasibility, and pivoting to get a new feasible solution, repeating until you find a solution.
- In the Simplex Method, if there are no positive coefficients corresponding to the nonbasic variables in the objective function, then you are at an optimal solution.
- In the Simplex Method, if there are no choices for the leaving variable, then the solution is unbounded.

Linear programming is a mathematical method for determining a way to achieve the best outcome for some list of requirements represented as linear relationships.

Example

A factory makes three types of chair, A, B and C. The factory makes a profit of $2 on A, $3 on B and $4 on C. A requires 30 man-hours, B requires 20 and C requires 10. A needs 2m2 of wood; B needs 5m2; C needs 3m2. Given 100 man-hours and 15m2 of wood per week, how many chairs of each type should be made each week to maximize profit?

Simplex Method

The most common method in linear programing is the simplex method. To use the simplex method, we need to represent the problem by linear equations. Let a be the number of A chairs, b the B chairs, and c the C chairs.

$30a + 20b + 10c \leq 100 \Rightarrow 3a + 2b + c \leq 10$

$2a + 5b + 3c \leq 15$

All three variables must be non-negative as well. All **constraints** must be satisfied.

The function to be maximized (the **objective function**) is:
$P = 2a + 3b + 4c$

Standard Form

The standard form for the simplex method is:

Minimize $c \cdot x$

Subject to: $Ax = b, x_i \geq 0$

Where $x = [x_1, x_2, \ldots, x_n]^T$ are the variables, $c = [c_1, c_2, \ldots, c_n]$ are the coefficients of the objective function, A is the left-side of the constraints, and $b = [b_1, b_2, \ldots, b_p]^T$ the right.

The solution of a linear program is accomplished in two steps. In the first step, Phase I, a starting extreme point is found. Phase I either gives a basic feasible solution or no solution. In the latter, the linear program is infeasible. In the second step, Phase II, the simplex algorithm is applied using the solution found in Phase I as a starting point. The possible results from Phase II are either an optimal solution or an unbounded solution.

Achieving Standard Form

You may have noticed that we had inequalities, $3a + 2b + c \leq 10$, but standard form calls for equalities. We therefore introduce a slack variable that represents the difference between the two sides of the inequality and is non-negative. This gives us the new equality $3a + 2b + c + s = 10$. The other inequality becomes: $2a + 5b + 3c + t = 15$.

Standard form also requires the objective function be a minimization. If the problem calls for maximization, multiply the objective function by -1.

Here are the pieces for standard form:

$x = [a, b, c, s, t]^T$

$c = [-2, -3, -4, 0, 0]$

$$A = \begin{bmatrix} 3 & 2 & 1 & 1 & 0 \\ 2 & 5 & 3 & 0 & 1 \end{bmatrix}$$

$$b = \begin{bmatrix} 10 \\ 15 \end{bmatrix}$$

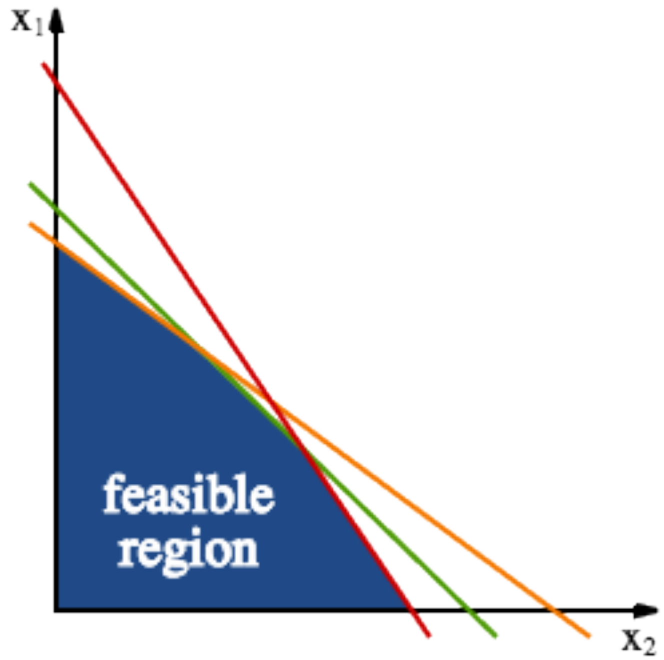
Canonical Tableaux

A linear program in standard form can be represented as a tableau of the form

$$\begin{bmatrix} 1 & -c & 0 \\ 0 & A & b \end{bmatrix}$$

The first row defines the objective function and the remaining rows specify the constraints. If the columns of A can be rearranged so that it contains the p-by-p identity matrix (the number of rows in A) then the tableau is said to be in canonical form. The variables corresponding to the columns of the identity matrix are called basic variables, while the remaining variables are called nonbasic or free variables. If the nonbasic variables are assumed to be 0, then the values of the basic variables are easily obtained as entries in b and this solution is a basic feasible solution.

Figure 6.23 Graph of Linear Inequality



The blue area is the feasible region for a linear program with the constraints forming the lines. The optimal solution will occur at one of the intersection points of the lines.

Pivots

Moving from one basic feasible solution to an adjacent basic feasible solution is called a pivot. First, a nonzero pivot element is selected in a nonbasic column. The row containing this element is multiplied by its reciprocal to change this element to 1, and then multiples of the row are added to the others to change the other entries in the column to 0. The result, is that if the pivot is in row r, then the column becomes the r-th column of the identity matrix. The variable for this column is now basic, replacing the variable which corresponded to the r-th column of the identity matrix. The variable corresponding to the pivot column enters the set of basic variables, and the variable being replaced leaves the set of basic variables.

Simplex Method Algorithm

The simplex algorithm proceeds by performing successive pivot operations which each improve the basic feasible solution; the choice of pivot element at each step is largely determined by the requirement that this pivot improves the solution.

For the entering variable, choose any column in which the entry in the objective row is positive. If all the entries in the objective row are less than or equal to 0 then no choice of entering variable can be made and the solution is optimal.

For the choice of pivot row, only positive entries in the pivot column are considered. This guarantees that the value of the entering variable will be non-negative. If there are none in the pivot column, then the entering variable can take any non-negative value with the

solution remaining feasible. Therefore the objective function is unbounded.

Next, the pivot row must be selected so that all the other basic variables remain positive. This occurs when the resulting value of the entering variable is at a minimum. If the pivot column is c, then the pivot row r is chosen so that b_r/a_{cr} is at a minimum.

Example

Using our example, the canonical tableau is

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 0 & 0 & 0 \\ 0 & 3 & 2 & 1 & 1 & 0 & 10 \\ 0 & 2 & 5 & 3 & 0 & 1 & 15 \end{bmatrix}$$

Columns 5 and 6 are the basic variables s and t, and the basic feasible solution is $a = b = c = 0, s = 10, t = 15$.

Columns 2, 3, and 4 can be selected as pivot columns; for this example column 4 is selected. The values of x resulting from the choice of rows 2 and 3 as pivot rows are 10/1 = 10 and 15/3 = 5 respectively. Of these the minimum is 5, so row 3 must be the pivot row. Performing the pivot produces:

$$\begin{bmatrix} 1 & \frac{-2}{3} & \frac{-11}{3} & 0 & 0 & \frac{-4}{3} & -20 \\ 0 & \frac{7}{3} & \frac{1}{3} & 0 & 1 & \frac{-1}{3} & 5 \\ 0 & \frac{2}{3} & \frac{5}{3} & 1 & 0 & \frac{1}{3} & 5 \end{bmatrix}$$

Now columns 4 and 5 represent the basic variables c and s and the corresponding basic feasible solution is

$a = b = t = 0, s = 5, c = 5$

For the next step, there are no positive entries in the objective row, and in fact: $-P = -20 + \frac{2}{3}a + \frac{11}{3}b + \frac{4}{3}t$

So we should make 5 chairs of type C to maximize our profits with 20 dollars.

Section 8

Partial Fractions

Partial Fractions

https://www.boundless.com/algebra/systems-of-equations-and-matrices/partial-fractions/

Partial Fractions

Partial fraction decomposition is a procedure used to reduce the degree of either the numerator or the denominator of a rational function.

KEY POINTS

- In terms of symbols, partial fraction decomposition turns a function of the form $\frac{f(x)}{g(x)}$, where f and g are both polynomials, into a function of the form $\sum_j \frac{f_j(x)}{g_j(x)}$, where gj(x) are polynomials that are factors of g(x).
- The main motivation to decompose a rational function into a sum of simpler fractions is to make it simpler to perform linear operations on the sum.
- If the degree of f(x) is greater than or equal to the degree of g(x), then it is necessary to perform the Euclidean division of f by g, using polynomial long division, giving f(x) = E(X)g(x) + h(x).
- If g(x) contains factors which are irreducible, then the numerator N(x) of each partial fraction with such a factor h(x) in the denominator must be sought as a polynomial with degree N < degree h, rather than as a constant.

$$R(x) = f(x)/g(x)$$
$$g(x) = P(x)*Q(x)$$
$$R(x) = A/P(x)+B/Q(x)$$

Figure 6.24 Partial Fraction Decomposition

This is the basic form of partial fraction decomposition.

Partial Fraction Decomposition

In algebra, the partial fraction decomposition or partial fraction expansion is a procedure used to reduce the degree of either the numerator or the denominator of a rational function (also known as a rational algebraic fraction).

In symbols, one can use partial fraction expansion to change a rational function in the form: $\frac{f(x)}{g(x)}$, where f and g are **polynomials**, into a function of the form, $\sum_j \frac{f_j(x)}{g_j(x)}$, where gj(x) are polynomials that are factors of g(x), and in general of lower degree. Thus, the partial fraction decomposition may be seen as the inverse procedure of the more elementary operation of addition of algebraic fractions, which produces a single rational function with a numerator and denominator usually of high degree. The full decomposition pushes the reduction as far as it will go: in other words, the factorization of g is used as much as possible.

The main motivation to decompose a rational function into a sum of simpler fractions is to make it simpler to perform linear operations on the sum. Therefore, the problem of computing derivatives, antiderivatives, integrals, power series expansions, Fourier series, Laplace transforms, residues, and linear functional transformations of rational functions can be reduced, via partial fraction decomposition, to focusing on computing each single element of the decomposition.

Basic Principles

Assume a rational function $R(x) = \frac{f(x)}{g(x)}$ has a denominator that factors as $g(x) = P(x) \cdot Q(x)$. If P and Q have no common factor, then R may be written as $\frac{A}{P} + \frac{B}{Q}$ for some polynomials A(x) and B(x).

The fraction can then be decomposed by equating $R(x) = \frac{A}{P} + \frac{B}{Q}$ and solving for each of A and B by substitution, equating the coefficients of terms, or otherwise.

There are some important cases to note for partial fraction decomposition.

If the degree of f(x) is greater than or equal to the degree of g(x), then it is necessary to perform the Euclidean division of f by g, using polynomial long division, giving f(x) = E(X)g(x) + h(x). Dividing through by g(x) gives $\frac{f(x)}{g(x)} = E(x) + \frac{h(x)}{g(x)}$, which you can then perform the decomposition on h(x)/g(x).

If g(x) contains factors which are irreducible, then the numerator N(x) of each partial fraction with such a factor h(x) in the denominator must be sought as a polynomial with degree N < degree h, rather than as a constant. For example, take the following decomposition over R:

$$\frac{x^2+1}{(x+2)(x-1)(x^2+x+1)} = \frac{a}{x-2} + \frac{b}{x-1} + \frac{cx+d}{x^2+x+1}$$

Example 1

$$f(x) = \frac{1}{x^2+2x-3}$$

Here, the denominator splits into two distinct linear factors:

$$q(x) = x^2 + 2x - 3 = (x+3)(x-1)$$

So we have the partial fraction decomposition:

$$f(x) = \frac{1}{x^2+2x-3} = \frac{A}{x+3} + \frac{B}{x-1}$$

Multiplying through by $x^2 + 2x - 3$, we have:

$1 = A(x-1) + B(x+3)$

Substitution x = -3 into this equation gives A = -1/4, and substituting x = 1 gives B = 1/4, so that:

$$f(x) = \frac{1}{x^2+2x-3} = \frac{1}{4}(\frac{-1}{x+3} + \frac{1}{x-1})$$

Example 2

$$f(x) = \frac{x^3+16}{x^3-4x^2+8x}$$

Because the numerator and denominator have the same degree, we have to perform long-divison. After long divison we get:

$$f(x) = 1 + \frac{4x^2-8x+16}{x^3-4x^2+8x} = 1 + \frac{4x^2-8x+16}{x(x^2-4x+8)}$$

x^2-4x+8 is irreducible, so we have:

$$\frac{4x^2-8x+16}{x(x^2-4x+8)} = \frac{A}{x} + \frac{Bx+C}{x^2-4x+8}$$

Multiply through by x^3-4x^2+8x, and we get:

$$4x^2-8x+16 = A(x^2-4x+8) + (Bx+C)x$$

Taking x = 0, we see that 16 = 8A, so A = 2. Comparing the x2 coefficients, we see that 4 = A + B = 2 + B, so B = 2. Comparing the linear coefficients, we see that -8 = -4A + C = -8 + C, so C = 0. Altogether,

$$f(x) = 1 + 2(\frac{1}{x} + \frac{x}{x^2-4x+8})$$

Source: https://www.boundless.com/algebra/concepts/6190/partial-fractions

Chapter 7

Conic Sections

https://www.boundless.com/algebra/conic-sections/

Section 1

The Parabola

https://www.boundless.com/algebra/conic-sections/the-parabola/

Parabolas

Parabolas are common in algebra as the graphs of quadratic functions, and they have many important real world applications.

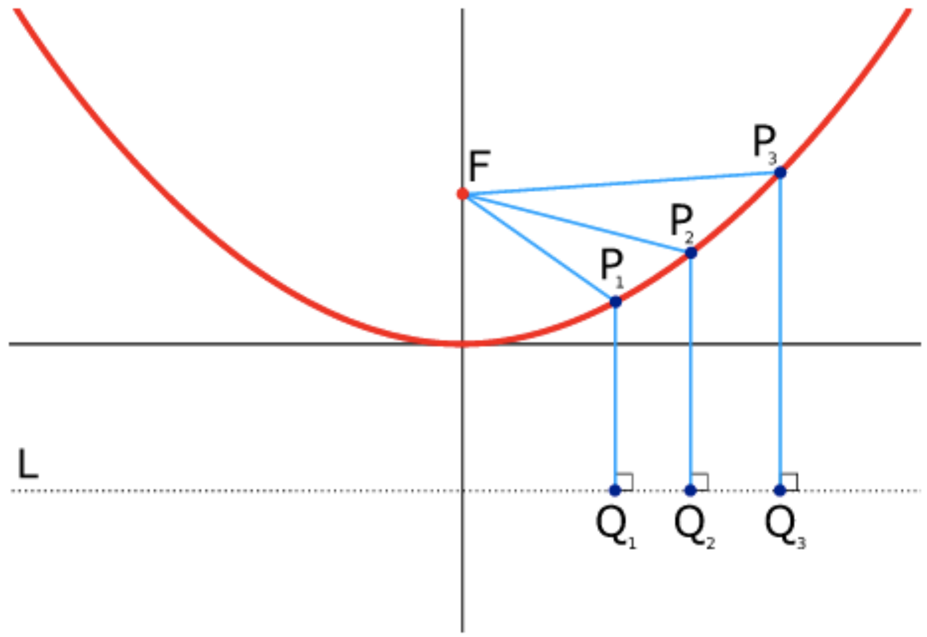
KEY POINTS

- Parabolas are frequently encountered as graphs of quadratic functions, including the very common equation y=x^2.
- All parabolas contain a focus, a directrix, and an axis of symmetry that vary in exact location depending on the equation used to define the parabola.
- Parabolas are frequently used in physics and engineering in places such as automobile headlight reflectors and in the design of ballistic missiles.

In mathematics, a parabola is a conic section, created from the intersection of a right circular conical surface and a plane parallel to a generating straight line of that surface. Another way to generate a parabola is to examine a point (the focus) and a line (the **directrix**), as can be visualized in *Figure 7.1*. The locus of points in that plane that are equidistant from both the line and point is a parabola. In algebra, parabolas are frequently encountered as graphs of quadratic functions, such as: $y = x^2$

The line perpendicular to the directrix and passing through the focus, that is, the line that splits the parabola through the middle, is called the axis of symmetry. The point on the axis of symmetry that intersects the parabola is called the "**vertex**", and it is the point where the curvature is greatest. The distance between the vertex and the focus, measured along the axis of symmetry, is the "focal length". Parabolas can open up, down, left, right, or in some other arbitrary direction. Any parabola can be repositioned and rescaled to fit exactly on any other parabola — that is, all parabolas are similar.

Figure 7.1 Parabola with Focus and Directrix



Parabolic curve showing directrix (L) and focus (F). The distance from any point on the parabola to the focus (PnF) equals the perpendicular distance from the same point on the parabola to the directrix (PnQn).

Parabolas have the property that, if they are made of material that reflects light, then light which enters a parabola traveling parallel to its axis of symmetry is reflected to its focus, regardless of where on the parabola the reflection occurs. Conversely, light that originates from a point source at the focus is reflected, or collimated, into a parallel beam, leaving the parabola parallel to the axis of symmetry. The same effects occur with sound and other forms of energy. This reflective property is the basis of many practical uses of parabolas.

The parabola has many important applications, from automobile headlight reflectors to the design of **ballistic** missiles. They are frequently used in physics, engineering, and many other areas.

Standard Form and Completing the Square

The standard form of a quadratic equation is useful for completing the square, which is used to graph the equation.

KEY POINTS

- Quadratic equations are graphed as parabolas.
- The standard form of a parabola is $y = a(x-h)^2 + k$.
- Completing the square is a technique used to solve for the x-intercepts, a necessary first step in graphing a quadratic equation as a parabola.

In algebra, **parabolas** are frequently encountered as graphs of quadratic functions, such as:

$$y = x^2$$

In algebraic geometry, the parabola is generalized by the rational normal curves, which have coordinates of $x,\ x^2,\ x^3,\ \ldots,\ x^n$. The standard parabola is the case n = 2. The case n = 3 is known as the twisted cubic. A further generalization is given by the Veronese variety, when there is more than one input variable.

Completing the square is a technique for converting a quadratic polynomial of the form:

$$y = ax^2 + bx + c$$

to the form:

$$y = a(\dots)^2 + constant$$

In this context, "constant" means not depending on x. The expression inside the parenthesis is of the form (x – constant). Thus one converts $ax^2 + bx + c$ to:

$$y = a(x - h)^2 + k$$

and one must find h and k. This form of quadratic equation is known as the "standard form" for graphing parabolas in algebra; from this equation, it is simple to determine the x-intercepts (y = 0) of the parabola, a process known as "solving" the quadratic equation.

Completing the square may be used to solve any quadratic equation. For example:

$$y = x^2 + 6x + 5$$

The first step is to set this equation equal to zero.

$$x^2 + 6x + 5 = 0$$

The second step is to complete the square:

$$(x + 3)^2 - 4 = 0$$

Next, we solve for the squared term:

$$(x + 3)^2 = 4$$

That gives us:

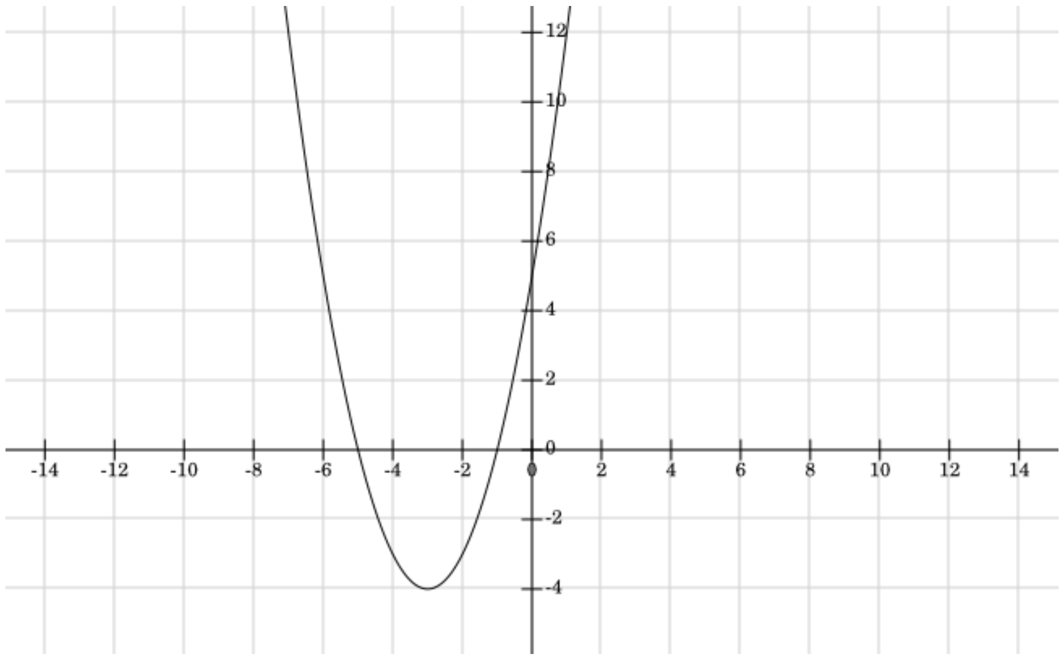
$x + 3 = -2$ or $x + 3 = 2$

And therefore:

$x = -5$ or $x = -1$

The result of graphing this equation can be seen in *Figure 7.2*, where the x-intercepts are indeed at -1 and -5. This can be applied to any quadratic equation. When the x2 has a coefficient other than 1, the first step is to divide out the equation by this coefficient.

Figure 7.2 Graph of y = x2 + 6x + 5



The graph of this quadratic equation is a parabola with x-intercepts at -1 and -5.

Source: https://www.boundless.com/algebra/conic-sections/the-parabola/standard-form-and-completing-the-square/

Applications and Problem Solving

Parabolas have important applications in physics, engineering, and nature.

KEY POINTS

- Projectiles and missiles follow an approximate parabolic path; they are approximate because real-world imperfections affect the movement of an object.
- Parabolic reflectors are common in today's microwave and satellite dish receiving and transmitting antennas.
- Paraboloids are also observed in the surface of a liquid confined to a container and rotated around the central axis.

A parabola is a conic section created from the intersection of a right circular **conical** surface and a plane parallel to a generating straight line of that surface. The parabola has many important applications, from automobile headlight reflectors to the design of ballistic missiles. They are frequently used in areas such as engineering, physics, and nature.

The parabolic trajectory of **projectiles** was discovered experimentally in the 17th century by Galileo, who performed

experiments with balls rolling on inclined planes. For objects extended in space, such as a diver jumping from a diving board, the object follows a complex motion as it rotates, while its center of mass forms a parabola.

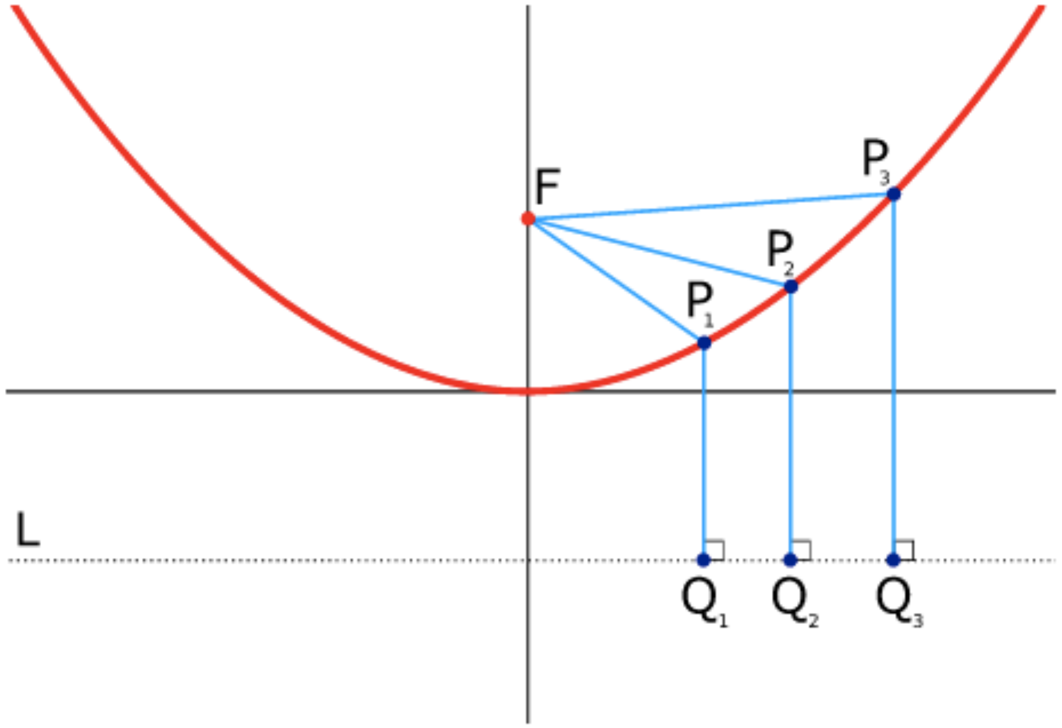
As in all cases in the physical world, a projectile's trajectory is an approximation. The presence of air resistance, for instance, distorts parabolic shape, although at low speeds the shape is a good approximation. At higher speeds, such as in **ballistics**, the shape is highly distorted and does not resemble a parabola (*Figure 7.3*).

Figure 7.3 Parabolic Water Trajectory

In this image, the water shot from a fountain follows a parabolic trajectory as gravity pulls it back down.

Paraboloids also arise in several physical situations. Best-known is the parabolic reflector—a mirror or similar reflective device that concentrates light or other forms of electromagnetic radiation to a common focal point, or conversely collimates light from a point source at the focus into a parallel beam (*Figure 7.4*). This principle was applied to telescopes in the 17th century. Today, paraboloid

Figure 7.4 Parabolic reflector



Note how the source light (Q1, Q2, Q3) reflects off different portions of the parabola, but collects at the same focal point (F).

reflectors are common throughout much of the world in microwave and satellite dish receiving and transmitting antennas.

Paraboloids are also observed in the surface of a liquid confined to a container then rotated around the central axis. In this case, he centrifugal force causes the liquid to climb the walls of the container, forming a parabolic surface (the principle behind the liquid mirror telescope).

Aircraft used to create a weightless state for purposes of experimentation, such as NASA's "Vomit Comet," follow a vertically

parabolic trajectory for brief periods in order to trace the course of an object in free fall, which, for most purposes, produces the same effect as zero gravity.

Section 2

The Circle and the Ellipse

https://www.boundless.com/algebra/conic-sections/the-circle-and-the-ellipse/

Circles

The set of all points in a plane that are the same distance from a given point forms a circle.

KEY POINTS

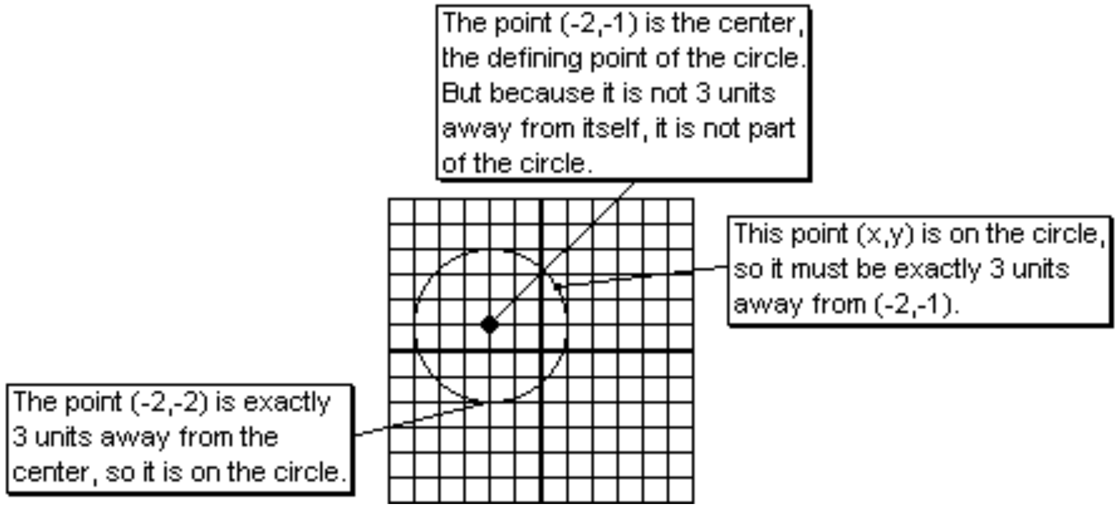
- The mathematical formula for a circle is (x–h)2+(y–k)2=r2, with the circle's center (h, k) and its radius, r.
- The length of a circumference of a circle can be found from the radius using the equation: $C = 2\pi r = \pi d$.
- The area enclosed by a circle can be found from the radius using the formula: $Area = \pi r^2$.

You've known all your life what a **circle** looks like. You probably know how to find the area and the circumference of a circle, given its radius. But what is the exact mathematical definition of a circle? Before you read the answer, you may want to think about the question for a minute. Try to think of a precise, specific definition of exactly what a circle is.

Definition of a Circle

The set of all points in a plane that are the same distance from a given point forms a circle. The point is known as the center of the circle, and the distance is known as the radius (*Figure 7.5*).

Figure 7.5 A Circle



A circle is defined as all the points that are a certain distance on a place from another point, the center.

Mathematicians often seem to be deliberately obscuring things by creating complicated definitions for things you already understood, but if you try to find a simpler definition of what a circle is, you will be surprised at how difficult it is to define. Most people start with something like "a shape that is completely round." That does describe a circle, but it also describes many other shapes, such as a pretzel.

So, you start adding caveats like "it can't cross itself" and "it can't have any loose ends." And then somebody draws an egg shape that fits all of your criteria, and yet is still not a circle. Hence, the definition for a circle as given above.

The Mathematical Formula for a Circle

You already know the formula for a line: y=mx+b. You know that m is the slope, and b is the y-intercept. Knowing all of this, you can easily answer questions such as: “Draw the graph of y=2x–3,” or “Find the equation of a line that contains the points (3,5) and (4,4).” If you are given the equation 3x+2y=6, you know how to graph it in two steps: first put it in the standard y=mx+b form, and then graph it.

All the conic sections are graphed in a similar way. There is a standard form which is very easy to graph, once you understand what all the parts mean. If you are given an equation that is not in standard form, you must put it into the standard form, and then graph it.

To understand the formula below, think of it as the y=mx+b of circles.

The mathematical formula for a circle is

$$(x-h)^2+(y-k)^2=r^2$$

with the circle's center (h, k) and its radius, r.

Length of Circumference

The ratio of a circle's circumference to its diameter is π (pi), an irrational constant approximately equal to 3.141592654. The length of the circumference, C, is related to the radius, r, and diameter, d, by:

$$C = 2\pi r = \pi d$$

Area Enclosed

As proved by Archimedes, the area enclosed by a circle is equal to that of a triangle whose base has the length of the circle's circumference, and whose height equals the circle's radius, which comes to π multiplied by the radius squared:

$$Area = \pi r^2$$

Equivalently, denoting diameter by d,

$$Area = \frac{\pi d^2}{4} \approx 0.7854 d^2,$$

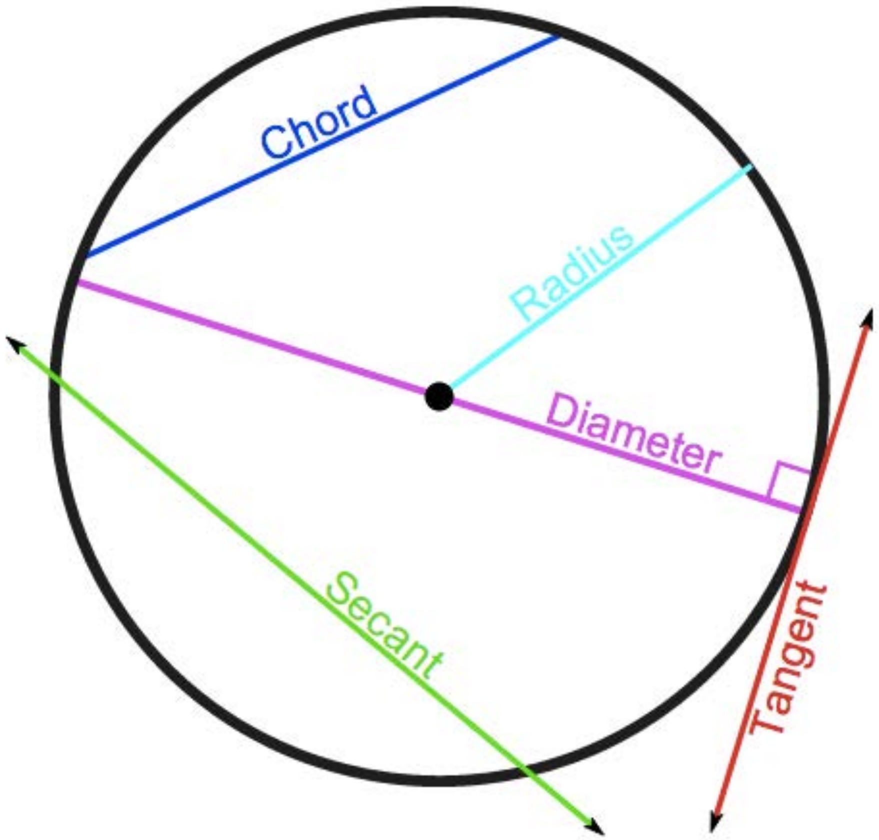
that is, approximately 79 percent of the circumscribing square (whose side is of length d).

Terminology of a Circle

- Chord: a line segment whose endpoints lie on the circle.

- Diameter: the longest chord, a line segment whose endpoints lie on the circle and which passes through the center; or the length of such a segment, which is the largest distance between any two points on the circle.
- Radius: a line segment joining the center of the circle to any point on the circle itself; or the length of such a segment, which is half a diameter.
- Circumference: the length of one circuit along the circle itself.

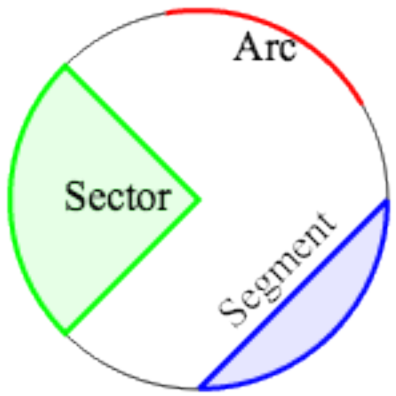
Figure 7.6 Circle terminology displayed



Chord, secant, tangent, radius, and diameter.

- Tangent: a straight line that touches the circle at a single point.
- Secant: an extended chord, a straight line cutting the circle at two points.
- Arc: any connected part of the circle's circumference.
- Sector: a region bounded by two radii and an arc lying between the radii.

Figure 7.7 Circle terminology



Arc, sector, and segment.

- Segment: a region bounded by a chord and an arc lying between the chord's endpoints.

See *Figure 7.6* and *Figure 7.7* for visual demonstrations.

Source: https://www.boundless.com/algebra/conic-sections/the-circle-and-the-ellipse/circles--2/

Ellipses

An ellipse, which resembles an oval, is defined as all points whose distance from two foci add to a constant.

KEY POINTS

- An ellipse is like a stretched out circle, sometimes referred to as an oval. In mathematics, an ellipse is a plane curve that results from the intersection of a cone by a plane in a way that produces a closed curve.
- An ellipse is also the locus of all points of the plane whose distances to two fixed points (each called a focus) add to the same constant.
- The formula for an ellipse is $\frac{(x-h)^2}{a^2}+\frac{(y-k)^2}{b^2}=1$. The area of an ellipse is given by πab, where a and b are one-half of the ellipse's major and minor axes respectively.

An **ellipse** is a sort of stretched out circle, sometimes referred to as an oval. In mathematics, an ellipse is a plane curve that results from the intersection of a cone by a plane in a way that produces a closed curve (*Figure 7.8*). Circles are special cases of ellipses, obtained when the cutting plane is orthogonal to the cone's axis. An ellipse has the property that, at any point on its perimeter, the distance from two fixed points (the foci) add to the same constant.

Drawing an Ellipse

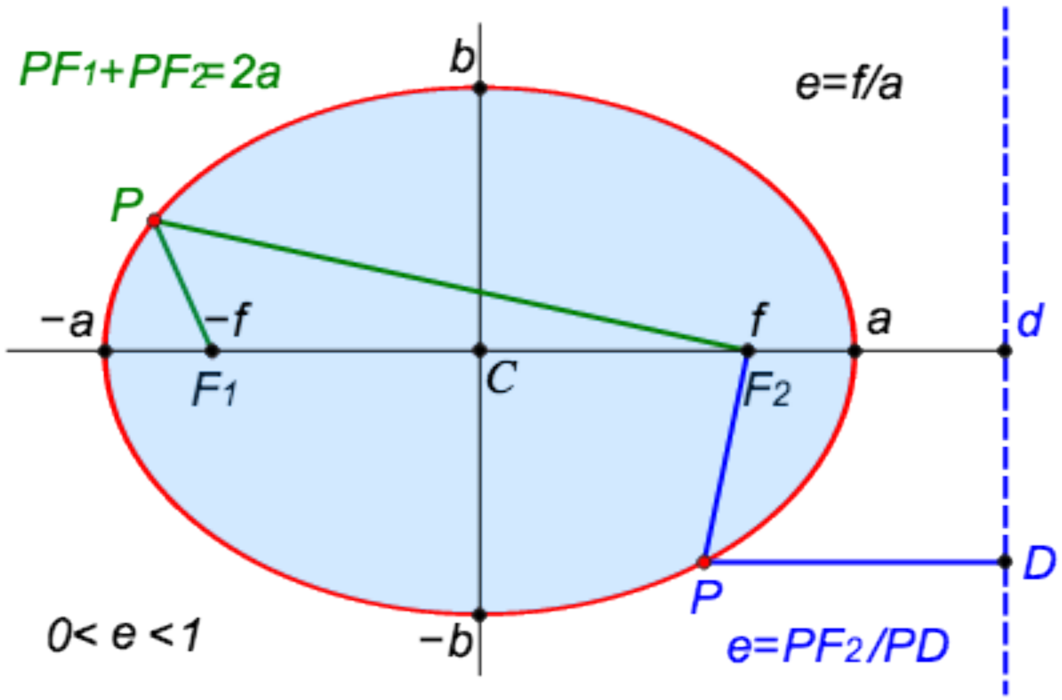
1. Lay a piece of cardboard on the floor.
2. Thumbtack one end of a string to the cardboard.
3. Thumbtack the other end of the string, elsewhere on the cardboard. The string should not be pulled taut: it should have some slack.
4. With your pen, pull the middle of the string back until it is taut.
5. Pull the pen all the way around the two thumbtacks, keeping the string taut at all times.
6. The pen will touch every point on the cardboard such that the distance to one thumbtack, plus the distance to the other thumbtack, is exactly one string length. And the resulting shape will be an ellipse. The cardboard is the "plane" in our definition, the thumbtacks are the "foci," and the string length is the "constant distance."

Figure 7.8 An ellipse as part of a cone

One definition of an ellipse is the intersection of a cone with an inclined plane.

How often do ellipses come up in real life? You'd be surprised. For a long time, the orbits of the planets were assumed to be circles. However, this is incorrect: the orbit of a planet is actually in the shape of an ellipse. The sun is at one focus of the ellipse (not at the center). Similarly, the moon travels in an ellipse, with the Earth at one focus.

Figure 7.9 Properties of an ellipse



See text for descriptions of each element of this ellipse.

Elements of an Ellipse

An ellipse is a smooth closed curve which is symmetric about its horizontal and vertical axes (*Figure 7.9*). The distance between antipodal points on the ellipse, or pairs of points whose midpoint is at the center of the ellipse, is maximum along the major axis or transverse diameter, and a minimum along the perpendicular minor axis or conjugate diameter.

The semi-major axis (denoted by a in the figure) and the semi-minor axis (denoted by b in the figure) are one half of the major and minor axes, respectively. These are sometimes called (especially in technical fields) the major and minor semi-axes, or major radius and minor radius. The four points where these axes cross the ellipse are the vertices, points where its curvature is minimized or maximized.

The foci of the ellipse are two special points F1 and F2 on the ellipse's major axis and are equidistant from the center point. The sum of the distances from any point P on the ellipse to those two foci is constant and equal to the major axis (PF1 + PF2 = 2a). Each of these two points is called a focus of the ellipse.

The eccentricity of an ellipse, usually denoted by ε or e, is the ratio of the distance between the two foci, to the length of the major axis or

$$e = \frac{2f}{2a} = \frac{f}{a}$$

For an ellipse the eccentricity is between 0 and 1.

The Formula of an Ellipse

With ellipses, it is crucial to start by distinguishing horizontal from vertical.

The general formula for an ellipse is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. If a>b, the ellipse is horizontal. If a, the the ellipse is vertical.

And of course, the usual rules of permutations apply. For instance, if we replace x with x–h, the ellipse moves to the right by h. So we have the more general form:

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$$

The area enclosed by an ellipse is πab, where a and b are one-half of the ellipse's major and minor axes respectively.

Source: https://www.boundless.com/algebra/conic-sections/the-circle-and-the-ellipse/ellipses/

Applications and Problem Solving

Circles and ellipses are encountered in everyday life, and knowing how to solve their equations is very important.

KEY POINTS

- The equation for a circle is $(x-h)^2 + (y-k)^2 = r^2$. Where r=radius; (h,k)= center point coordinates.
- The equation for an ellipse is $\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$. Where (h,k) = center point coordinates; a^2 = major/minor axis length; b^2 = major/minor axis length. Out of a and b, which ever value is greater, that will be the major axis.
- In order to determine the parameters of a circle or an ellipse, you must first put the equation into standard form as shown above.

Circles are all around you in everyday life, from tires on cars to buttons on coats, as well as on the tops of bowls, glasses and water bottles. Ellipses are less common, mostly encountered as the orbits of planets, but you should be able to find the area of a circle or an ellipse, or the circumference of a circle, based on information given to you in a problem. Some sample problems are shown below. The

answers will be given at the end of the atom, try not to look until you have attempted the problem

Exercise 1

Let's say you are a gardener, and you have just planted a lot of flowers that you want to water. The flower bed is 15 square feet. You are using a circular sprinkler system, and the water reaches 6 feet out from the center. The sprinkler is located, from the bottom left corner of the bed, 7 feet up, and 6 feet over.

1. If the flower bed was a graph with the bottom left corner being the origin, what would the equation of the circle be?
2. What is the area being watered by the sprinkler?
3. What percentage of the garden that is being watered?

Exercise 2

Now, let's take it the other way. $(x-4)^2+(y+8)^2=49$ is the equation for a circle.

a. What are the coordinates of the center of the circle?

b. What is the radius?

c. Draw the circle.

d. Find two points on the circle (by looking at your drawing) and plug them into the equation to make sure they work. (Show your work!)

Exercise 3

a. Put $2x^2+2y^2+8x+24y+60=0$ into standard canonical circle form, $(x-h)^2+(y-k)^2=r^2$.

b. What are the center and radius of the circle?

c. Draw the circle.

d. Find two points on the circle (by looking at your drawing) and plug them into the original equation to make sure they work.

Exercise 4

$$\frac{4x^2}{9}+25y^2=1$$

This sort of looks like an ellipse in standard form, doesn't it? It even has a 1 on the right. But it isn't. Because we have no room in our standard form for that 4 and that 25—for numbers multiplied by the x^2 and y^2 terms. How can we get rid of them, to get into standard form, while retaining the 1 on the right?

a. Rewrite the left-hand term, $\frac{4x^2}{9}$, by dividing the top and bottom of the fraction by 4. Leave the bottom as a fraction; don't make it a decimal.

b. Rewrite the right-hand term, 25y², by dividing the top and bottom of the fraction by 25. Leave the bottom as a fraction; don't make it a decimal.

c. Now, you're in standard form. what is the center?

d. How long is the major axis?

e. How long is the minor axis?

f. Graph it.

Answers

Exercise 1:

1. $(x-7)^2 + (y-6)^2 = 36$
2. $A_{sprinkler} = \pi * r^2 \quad A_{sprinkler} = \pi * 6^2 \quad A_{sprinkler} = 36\pi$
3. $Percentage_{watered} = \frac{A_{sprinkler}}{A_{flowerbed}} * 100\,\%$

   $Percentage_{watered} = \frac{36\pi}{15^2} * 100\,\%$

   $Percentage_{watered} = \frac{113.1}{225} * 100\,\%$
4. $Percentage_{watered} = 50.3\,\%$

Answer 2:

a. (4, -8).

b. $\sqrt{49} = 7$

c. *Figure 7.10*

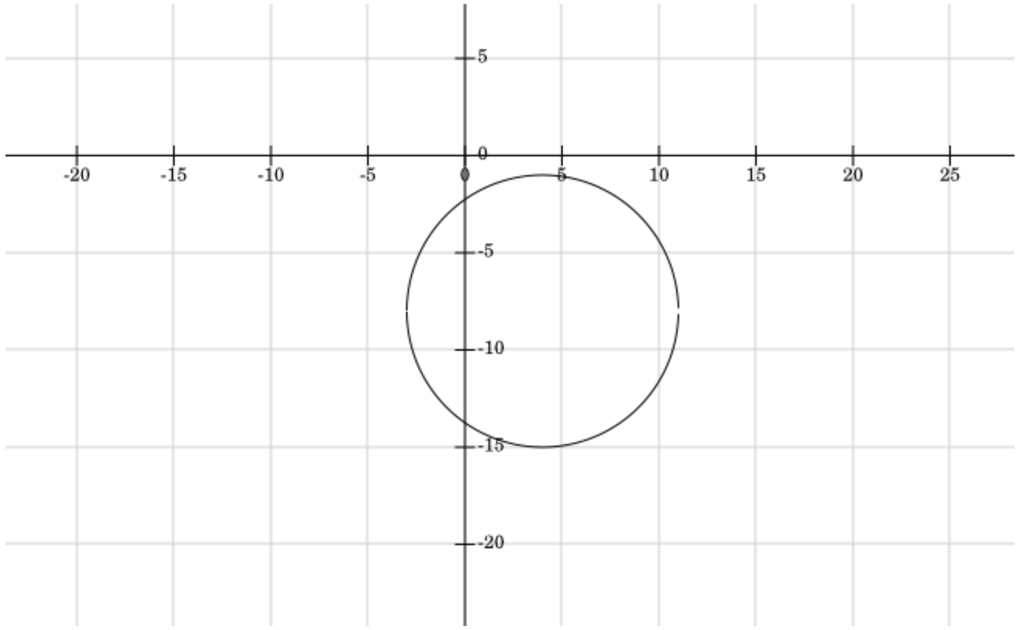
d. 4, -1 is the top of the circle, looking at the graph. plugging this into the equation, we get $(4-4)^2 + (-1+8)^2 = 49$

$0^2 + (-7)^2 = 49$

$49 = 49$

Therefore this is a valid point on the circle.

Figure 7.10 A graph of (x-4)^2+(y+8)^2=49 gives a circle



In order to graph this circle, the program required input in terms of y=, which is not the form for the circle. In order to graph it, we rearrange the equation for a circle to two equations, y=-8-sqrt(33+8 x-x^2) and y=-8+sqrt(33+8 x-x^2) and ignore the imaginary sections.

Answer 3:

a. First, divide by the coefficient of x^2 and y^2 to give
$x^2 + y^2 + 4x + 12y + 30 = 0$

Next, collect x and y terms together, and bring the number to the other side to give $(x^2 + 4x) + (y^2 + 12y) = -30$

Now, complete the square in both parentheses, subtracting or adding the necessary constant to both sides,
$(x^2 + 4x + 4) + (y^2 + 12y + 36) = -30 + 4 + 36$

Notice that each term is a perfect square, which gives
$(x + 2)^2 + (y + 6)^2 = 10$

b. The center is at (-2,-6) and the radius is $\sqrt{10}$

c. and d. are left up to the student.

Answer 4:

a. After dividing the top and bottom by 4, we get $\frac{x^2}{\frac{9}{4}} + 25y^2 = 1$

b. If we divide the top and bottom of the y term (in fraction form) by 25, we get $\frac{x^2}{\frac{9}{4}} + \frac{y^2}{\frac{1}{25}} = 1$. Now the equation is in our standard form.

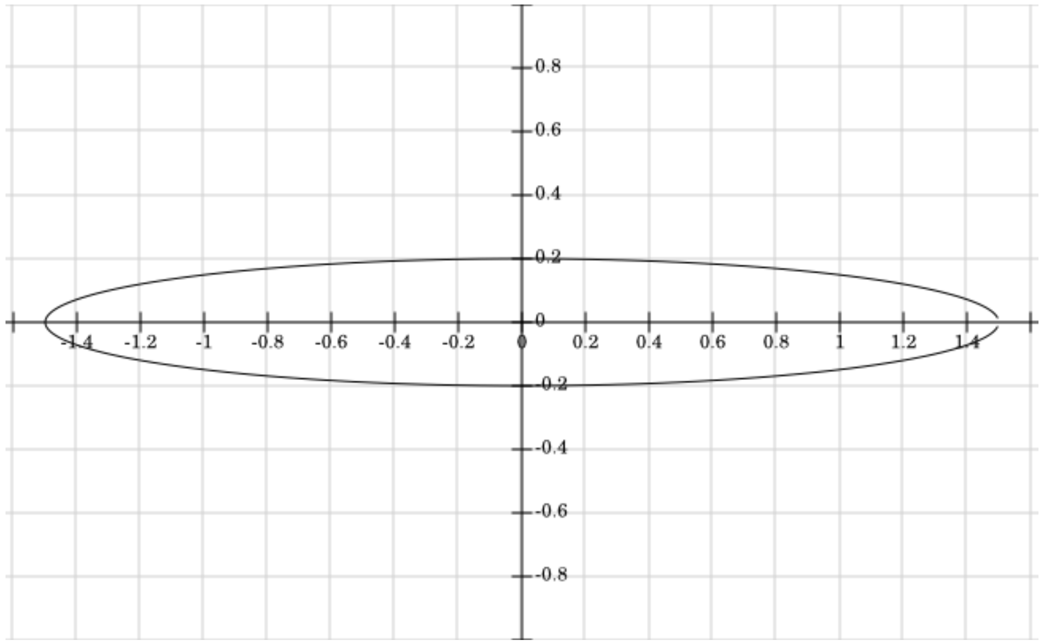
c. From our standard equation ($\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1$), we know that the center is at (h, k). Since these are both zero in our equation, the center is at (0,0).

d. The major axis is 2a, or $\frac{18}{4}$.

e. The minor axis is 2b, or $\frac{2}{25}$.

f. *Figure 7.11*

Figure 7.11 The ellipse given by x^2/(9/4)+y^2/(1/25)=1



In order to graph this ellipse, we must rearrange the equation in terms of y= (unless you have a calculator that can handle ellipses), so we graph + and - sqrt(9-4x^2)/(15)

Source: https://www.boundless.com/algebra/conic-sections/the-circle-and-the-ellipse/applications-and-problem-solving--6/

Section 3

The Hyperbola

Standard Equations of Hyperbolas

Applications and Problem Solving

https://www.boundless.com/algebra/conic-sections/the-hyperbola/

Standard Equations of Hyperbolas

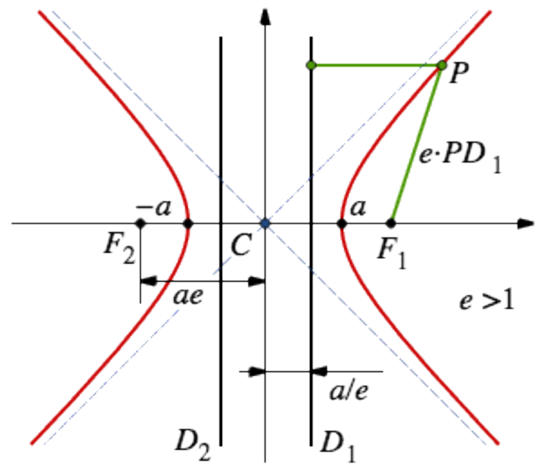
A standard equation for a hyperbola can be written as $x^2/a^2 - y^2/b^2 = 1$.

KEY POINTS

- Similar to a parabola, a hyperbola is an open curve, meaning that it continues indefinitely to infinity, rather than closing on itself as an ellipse does. A hyperbola consists of two disconnected curves called its arms or branches.
- If the transverse axis of any hyperbola is aligned with the x-axis of a Cartesian coordinate system and is centered on the origin, the equation of the hyperbola can be written as x^2/a^2 - $y^2/b^2 = 1$.
- Likewise, a hyperbola with its transverse axis aligned with the y-axis is called a "North-South opening hyperbola" and has equation $y^2/a^2 - x^2/b^2 = 1$.

Similar to a **parabola**, a **hyperbola** is an open curve, meaning that it continues indefinitely to infinity, rather than closing on itself as an ellipse does. A hyperbola consists of two disconnected curves called its arms or branches.

Figure 7.12 Properties of a Hyperbola



The asymptotes of the hyperbola (red curves) are shown as blue dashed lines and intersect at the center of the hyperbola, C. The two focal points are labeled F1 and F2, and the thin black line joining them is the transverse axis. The perpendicular thin black line through the center is the conjugate axis. The two thick black lines parallel to the conjugate axis (thus, perpendicular to the transverse axis) are the two directrices, D1 and D2. The eccentricity e equals the ratio of the distances from a point P on the hyperbola to one focus and its corresponding directrix line (shown in green).

At large distances from the center, the hyperbola approaches two lines, its **asymptotes**, which intersect at the hyperbola's center. A hyperbola approaches its asymptotes arbitrarily closely as the distance from its center increases, but it never intersects them. Consistent with the symmetry of the hyperbola, if the transverse axis is aligned with the x-axis, the slopes of the asymptotes are equal in magnitude but opposite in sign, ±b⁄a, where b=a×tan(θ) and where θ is the angle between the transverse axis and either asymptote. The distance b (not shown in *Figure 7.12*) is the length of the perpendicular segment from either vertex to the asymptotes.

A conjugate axis of length 2b, corresponding to the minor axis of an ellipse, is sometimes drawn on the non-transverse principal axis; its

endpoints ±b lie on the minor axis at the height of the asymptotes over/under the hyperbola's vertices. Because of the minus sign in some of the formulas below, it is also called the imaginary axis of the hyperbola.

If b = a, the angle 2θ between the asymptotes equals 90° and the hyperbola is said to be rectangular or equilateral. In this special case, the rectangle joining the four points on the asymptotes directly above and below the vertices is a square, since the lengths of its sides 2a = 2b.

If the transverse axis of any hyperbola is aligned with the x-axis of a Cartesian coordinate system and is centered on the origin, the equation of the hyperbola can be written as:

$$x^2/a^2 - y^2/b^2 = 1$$

A hyperbola aligned in this way is called an "East-West opening hyperbola." Likewise, a hyperbola with its transverse axis aligned with the y-axis is called a "North-South opening hyperbola" and has equation:

$$y^2/a^2 - x^2/b^2 = 1$$

Applications and Problem Solving

A hyperbola is an open curve with two branches and a cut through both halves of a double cone, which is not necessarily parallel to the cone's axis.

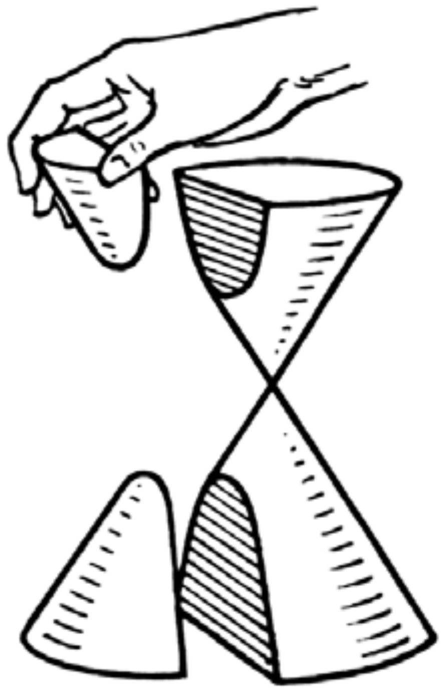
KEY POINTS

- Hyperbolas have applications to a number of different systems and problems including sundials and trilateration.
- Hyperbolas may be seen in many sundials. On any given day, the sun revolves in a circle on the celestial sphere, and its rays striking the point on a sundial traces out a cone of light. The intersection of this cone with the horizontal plane of the ground forms a conic section.
- A hyperbola is the basis for solving trilateration problems, the task of locating a point from the differences in its distances to given points — or, equivalently, the difference in arrival times of synchronized signals between the point and the given points.

As we should know by now, a hyperbola is an open curve with two branches, the intersection of a plane with both halves of a double cone. The plane may or may not be parallel to the axis of the cone

Sundials

Figure 7.13 Hyperbola



A hyperbola is an open curve with two branches, the intersection of a plane with both halves of a double cone. The plane may or may not be parallel to the axis of the cone

Hyperbolas may be seen in many sundials. Every day, the sun revolves in a circle on the celestial sphere, and its rays striking the point on a sundial traces out a cone of light. The intersection of this cone with the horizontal plane of the ground forms a **conic section**. At most populated latitudes and at most times of the year, this conic section is a hyperbola. This conic section can be shown in *Figure 7.13*. The shadow of the tip of a pole traces out a hyperbola on the ground over the course of a day (this path is called the declination line). The shape of this hyperbola varies with the geographical latitude and with the time of the year, since those factors affect the cone of the sun's rays relative to the horizon (*Figure 7.14*).

Trilateration

Trilateration is the a method of pinpointing an exact location, using its distances to a given points. The can also be characterized

as the

Figure 7.14 Hyperbolas and Sundials

Hyperbolas as declination lines on a sundial.

difference in arrival times of synchronized signals between the desired point and known points.These types of problems arise in navigation, mainly nautical. A ship can locate its position using the arrival times of signals from GPS transmitters. Alternatively, a homing beacon can be located by comparing the arrival times of its signals at two separate receiving stations. This can be used to track people, cell phones, internet signals and many other things.In particular, the set of possible positions of a point that has a distance variation of 2a from two known points is a hyperbola of vertex separation 2a, and whose foci are the two known points.

The Kepler Orbit of Particles

The Kepler orbit is the path followed by any orbiting body (*Figure 7.15*). This can be applied to a particle of any size, a planet or even hydrogen atoms. depending on the particles properties, including size and shape (eccentricity), this orbit can be one of six conic sections. In particular, if the total energy E of the particle is greater than zero (i.e., if the particle is unbound), the path of such a particle is a hyperbola. In the figure, the blue line shows the hyperbolic Kepler orbit.

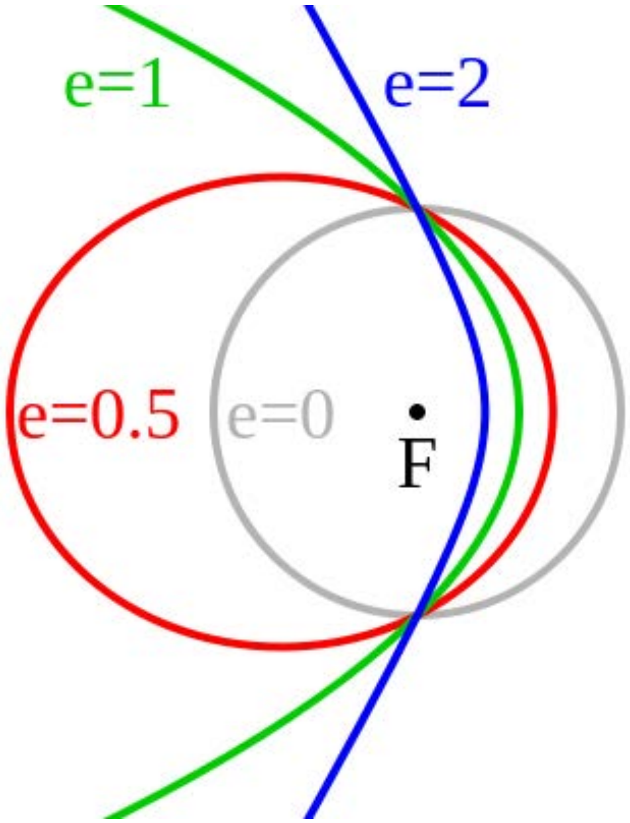


Figure 7.15 Kepler Orbits

A diagram of the various forms of the Kepler Orbit and their eccentricities. Blue is a hyperbolic trajectory ($e > 1$). Green is a parabolic trajectory ($e = 1$). Red is an elliptical orbit ($e < 1$). Grey is a circular orbit ($e = 0$).

EXAMPLE

A hyperbola is the basis for solving trilateration problems, the task of locating a point from the differences in its distances to given points — or, equivalently, the difference in arrival times of synchronized signals between the point and the given points. Such problems are important in navigation, particularly on water; a ship can locate its position from the difference in arrival times of signals from GPS transmitters.

Section 4

Nonlinear Systems of Equations and Inequalities

https://www.boundless.com/algebra/conic-sections/nonlinear-systems-of-equations-and-inequalities/

Nonlinear Systems of Equations and Problem Solving

As with linear systems, a nonlinear system of equations (and conics) can be solved graphically and algebraically for all its variables.

KEY POINTS

- Subtracting one equation from another is an effective means for solving linear systems, but it often is difficult to use in nonlinear systems, in which the terms of two equations may be very different.
- Substitution of a variable into another equation is usually the best method for solving nonlinear systems of equations.
- Nonlinear systems of equations may have one or multiple solutions.

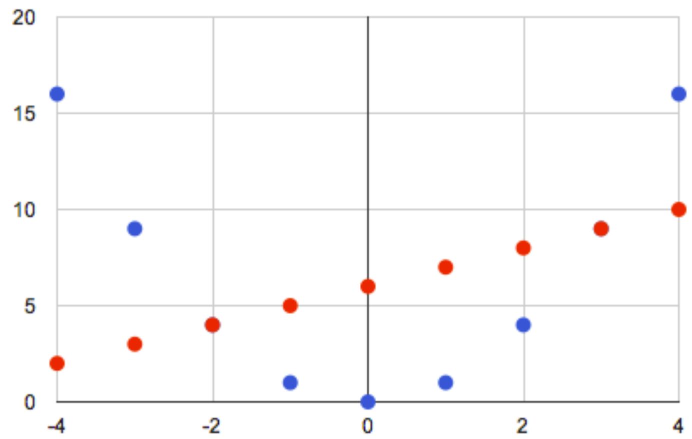
A **conic section** (or just conic) is a curve obtained as the intersection of a cone (more precisely, a right circular conical surface) with a plane. In analytic geometry, a conic may be defined as a plane algebraic curve of degree 2. There are a number of other geometric definitions possible. Traditionally, the three types of conic section are the hyperbola, the parabola, and the ellipse. The circle is a special case of the ellipse, and is of sufficient interest in its own right that it is sometimes called the fourth type of conic section. The type of a conic corresponds to its eccentricity, those with eccentricity less than 1 being ellipses, those with eccentricity equal to 1 being parabolas, and those with eccentricity greater than 1 being hyperbolas. In the focus-directrix definition of a conic, the circle is a limiting case with eccentricity 0. In modern geometry, certain degenerate cases, such as the union of two lines, are included as conics as well.

In a **system of equations**, two or more relationships are stated among variables. A system is solvable so long as there are as many simultaneous equations as variables. If each equation is graphed, the solution for the system can be found at the point where all the functions meet. The solution can be found either by inspection of a graph, typically with the use of software, or algebraically.

Nonlinear systems of equations, such as conic sections, include at least one function that is non-linear. Because at least one function has curvature, it is possible for nonlinear systems of equations to contain multiple solutions. As with linear systems of equations, substitution can be used to solve nonlinear systems for one variable and then the other.

Solving nonlinear systems of equations algebraically is similar to doing the same for linear systems of equations. However, subtraction of one equation from another can become impractical if the two equations have different terms, which is more commonly the case in nonlinear systems.

Figure 7.16 Integer values of y=x^2 (blue) and y=x+6 (red)



The parabola (blue) falls below the line (red) between x=-2 and x=3. For all values of x less than -2 and greater than 3, the parabola is greater than the line.

Consider, for example, the following system of equations (*Figure 7.16*):

$y = x^2$ (1)

$y = x + 6$ (2)

Substituting x2 for y in equation 2:

$x^2 = x + 6$

This quadratic equation can be solved by moving all the equation's components to the left before using the quadratic formula:

$x^2 - x - 6 = 0$

Using the quadratic formula, with a=1, b=-2 and c=-6, it can be determined that x=3 and x=-2 are solutions.

The solutions for x can then be plugged into either of the original systems to find the value of y. In this example, we will use equation 1:

$y = (-2)^2$

$y = 3^2$

Thus, for x=-2, y=4. And for x=3, y=9.

Our final solutions are: (-2, 4) and (3, 9).

Models Involving Nonlinear Systems of Equations

Nonlinear systems of equations can be used to solve complex problems involving multiple known relationships.

KEY POINTS

- Problems involving simultaneously moving bodies can be solved using systems of equations. If at least one body accelerates or decelerates, the system is nonlinear.
- If the relationship between multiple unknown numbers is described in as many ways as there are numbers, all can be found using systems of equations. If at least one of those relationships is nonlinear, the system is nonlinear.
- Substitution is the best method for solving for simultaneous equations, although to answer a question, one may not need to solve for every variable.

Nonlinear **systems of equations** are not just for hypothetical discussions; they can be used to solve complex problems involving multiple known relationships.

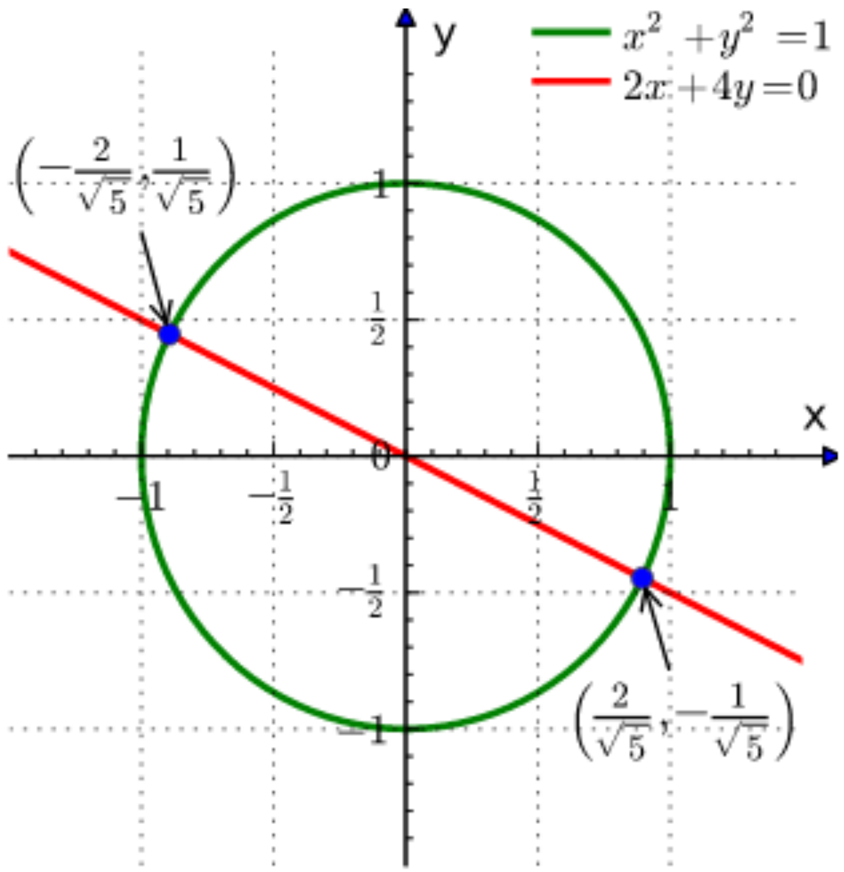
Consider, for example, a car that begins at rest and accelerates at a constant rate of 4 meters per second each second. Its position in meters (y) can be determined as a function of time in seconds (t), by the formula:

$y = 2t^2$ (1)

Now consider a second car, traveling at a constant speed of 20 meters per second. Its position (y) in meters can be determined as a function of time (t) in seconds, using the following formula:

$y = 20t$ (2)

Figure 7.17 Nonlinear Simultaneous Equations



When the first car begins to accelerate, the second car is 400 meters ahead of it. To express the position of the second car relative to the first as a function of time, we can modify the second equation as such:

$y = 20t + 400$ (3)

To determine where the cars are when they are alongside one another and how much time has passed since the first began to accelerate, we can algebraically solve the system of equations using substitution:

$2t^2 = 20t + 400$ (4)

Solving for t, we can find that the cars are side-by-side after 20 seconds.

Substituting 20 for t in either equation 2 or 3, we can find that the cars meet 800 meters ahead of the first car's starting point. Note that a question on an exam may not prompt solutions for both variables.

In addition to practical scenarios like the above, nonlinear systems can be alluded to in word problems. For example, a question could ask:

The product of two numbers is 12, and the sum of their squares is 40. What are the numbers?

In this case, we could make an equation for each known relationship:

$x * y = 12$ (5)

$x^2 + y^2 = 40$ (6)

Substitution can be used to prove that the numbers are 2 and 6.

Nonlinear Systems of Inequalities

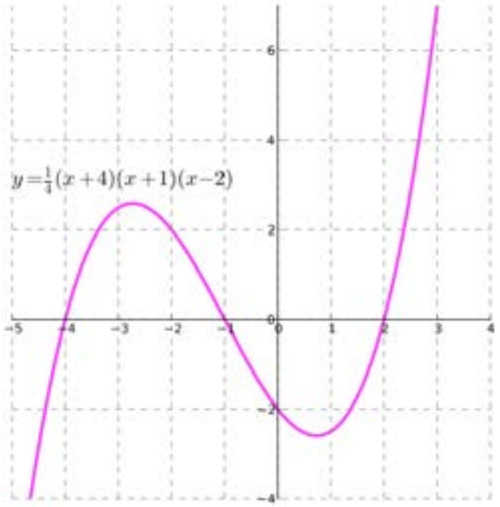
A nonlinear inequality contains two expressions connected by a "greater than" or "less than" signs and involves a nonlinear function.

KEY POINTS

- A nonlinear system of inequalities may have at least one solution; if it does, a solution may be bounded or unbounded.
- A solution for a nonlinear system of inequalities will be in a region that satisfies every inequality in the system.
- The best way to show solutions to nonlinear systems of inequalities is graphically, shading the area that satisfies all the system's constituent inequalities.

A system of inequalities consists of two expressions connected by a "greater than" (>) or "less than" (<) sign. A nonlinear inequality is an inequality that involves a nonlinear function - a polynomial function of degree 2 or higher, as shown in *Figure 7.18*. When operating in terms of real numbers, nonlinear inequalities are written in the forms $f(x) < b$ or $f(x) \leq b$ where f(x) is a nonlinear function and b is a constant real number.

Figure 7.18 The graph of a polynomial function of degree 3



Consider, for example, the system including the parabolic nonlinear inequality:

$$y > x^2$$

and the linear inequality:

$$y < x + 2$$

All points below the line y=x+2 satisfy the linear equality, and all points above the parabola y=x2 satisfy the parabolic nonlinear inequality.

Graphing both inequalities reveals one region of overlap: The area where the parabola dips below the line. This area is the solution to the system.

The limits of each inequality intersect at (-1, 1) and (2, 4). Note that the area above y=x2 that is also below y=x+2 is closed between those two points. Whereas a solution for a linear system of equations will contain an infinite, unbounded area (lines can only pass one another a maximum of once), in many instances, a solution for a nonlinear system of equations will consist of a finite, bounded area.

This need not be the case with all nonlinear inequalities, however: reversing the direction of both inequalities in the previous example would lead to an infinite solution area.

The best way to write solutions to systems of inequalities, both linear and nonlinear, is graphically. Each inequality can be drawn as a solid (if $<$ or $>$ is used) or dashed (if $\leq$ or $\geq$ is used) line. The region that satisfies all inequalities can be indicated with shading.

Chapter 8

Sequences, Series and Combinatorics

https://www.boundless.com/algebra/sequences-series-and-combinatorics/

Section 1

Sequences and Series

https://www.boundless.com/algebra/sequences-series-and-combinatorics/sequences-and-series/

Introduction to Sequences

In mathematics, a sequence is an ordered list of objects, often numbers and often defined in terms of the previous member of the set.

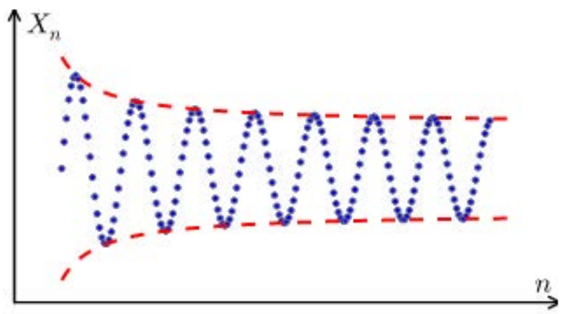
KEY POINTS

- The number of ordered elements (possibly infinite) is called the length of the sequence. Unlike a set, order matters, and exactly the same elements can appear multiple times at different positions in the sequence.
- An arithmetic sequence is arrived at by adding a constant to the previous term of a sequence to arrive at the next term. It can be described by the formula $a_n = a_m + (n - m)d$.
- A geometric sequence is one in which the previous member of a sequence is multiplied by a constant to arrive at the next term. It can be described by the formula $t_n = r \cdot t_{n-1}$.

In mathematics, a **sequence** is an ordered list of objects (or events). Like a **set**, it contains members (also called elements or terms). The number of ordered elements (possibly infinite) is called the length of the sequence. Unlike a set, order matters, and exactly the same elements can appear multiple times at different positions in the sequence. A sequence is a discrete function.

For example, (M, A, R, Y) is a sequence of letters that differs from (A, R, M, Y), as the ordering matters, and (1, 1, 2, 3, 5, 8), which contains the number 1 at two different positions, is a valid sequence. Sequences can be **finite**, as in this example, or infinite, such as the sequence of all even positive integers (2, 4, 6,...). Finite sequences are sometimes known as strings or words and infinite sequences as streams. The empty sequence () is included in most notions of sequence, but may be excluded depending on the context (*Figure 8.1*).

Figure 8.1 Sequence



An infinite sequence of real numbers (in blue). This sequence is neither increasing, nor decreasing, nor convergent. It is, however, bounded.

Examples and Notation: Finite and Infinite

A more formal definition of a finite sequence with terms in a set S is a function from {1, 2,..., n} to S for some n > 0. An infinite sequence in S is a function from {1, 2,... } to S. For example, the sequence of prime numbers (2,3,5,7,11, ...) is the function 1→2, 2→3, 3→5, 4→7, 5→11, ...

A sequence of a finite length n is also called an n-tuple. Finite sequences include the empty sequence () that has no elements.

A function from all integers into a set is sometimes called a bi-infinite sequence or two-way infinite sequence. An example is the bi-infinite sequence of all even integers (..., -4, -2, 0, 2, 4, 6, 8...).

Next, we will look at arithmetic, quadratic, and geometric sequences.

Arithmetic and Geometric Sequences

Many of the sequences you will encounter in a mathematics course are produced by a formula, where some operation(s) is performed on the previous member of the sequence (a_{n-1}) to give the next member of the sequence (a_n).

An arithmetic (or linear) sequence is a sequence of numbers in which each new term is calculated by adding a constant value to the previous term. An example is 10,13,16,19,22,25. In this example, the first term (which we shall call t_1 in this case, though a_1 would be equally valid) is 10, and the common difference (d)—that is, the difference between any two adjacent numbers—is 3. Another example is 25,22,19,16,13,10. In this example t_1=25, and d=(–3). In both of these examples, n (the number of terms) is 6.

A geometric sequence is a list where each number is generated by multiplying a constant by the previous number. An example is 2,6,18,54,162. In this example, t_1=2, and the common ratio (r)—that is, the ratio between any two adjacent numbers—is 3. Another example is 162,54,18,6,2. In this example t_1=162, and r=$\frac{1}{3}$. In both examples n=5.

A recursive definition of a sequence means that you define each term based on the previous. So the recursive definition of an arithmetic sequence is $t_n=t_{n-1}+d$, and the recursive definition of a geometric sequence is $t_n=rt_{n-1}$.

An explicit definition of an arithmetic sequence means you define the nth term without making reference to the previous term. This is more useful, because it means you can find (for instance) the 20th term without finding all the other terms in between.

To find the explicit definition of an arithmetic sequence, you just start writing out the terms. The first term is always t1. The second term goes up by d so it is t_1+d. The third term goes up by d again, so it is (t_1+d)+d, or in other words, t_1+2d. So we get a chart like this: $t_1=t_1$, $t_2=t_1+d$, $t_3=t_1+2d$, $t_2=t_1+3d$,... and so on. From this you can see the generalization that $t_n=t_1+(n-1)d$, which is the explicit definition we were looking for.

The explicit definition of a geometric sequence is arrived at the same way. The first term is t_1; the second term is r times that, or t_1r; the third term is r times that, or t_1r^2; and so on. So the general rule is $t_n = (t_1 r)^{n-1}$. Read this as: t1 multiplied by r, (n–1) times.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/sequences-and-series/introduction-to-sequences/

Finding the General Term

Given terms in a sequence, there are many ways to find the term based on variables that, when replaced with integers, affords valid terms.

KEY POINTS

- Plotting points on a Cartesian coordinate plane and using graphing software to solve is the easiest and most effective way to find a general term. Given terms $x_1, x_2 \ldots x_n$, convert these to points (1, x_1), (2, x_2)... (n, x_n) and plot. Then find the best-fit curve.
- By hand, one can take the differences between each term, then the differences between the differences in terms, etc, until each difference becomes constant.
- Once a constant difference is achieved, one can work backwards to find the relationship among terms in each set of the differences. Eventually, one can work backwards to obtain a sum for the initial series.

Given several values in a **sequence** or **series**, there are a number of ways to determine the **general term**. The general term is an expression that consists of variables and constants that, when substituting integers in place of x, produces a valid term in a sequence or series.

Find Differences Among Terms

In some instances, a sequence or series is simple enough to be solved generally by hand. One quick way is to take the differences between each term, then the differences between the differences in terms, etc, until each difference becomes constant. From there, one can work backwards to find the relationship among terms in each set of the differences. Eventually, one can work backwards to obtain a sum for the initial series.

For example, consider the sequence:

4, -7, -26, -53, -88, -131

The difference between 4 and -7 is -11; the difference between -7 and -26 is -19. Finding all these differences, we get a new set:

-11, -19, -27, -35, -43

This set is still not yet constant. However, finding the differences between terms once more, we get:

-8, -8, -8, -8

The formula for the last row of terms (f(n)) as a function of number of the term (n) is a simple constant:

$$f(n) = -8$$

We can use this function in relating the previous row of terms (-11, -19, -27, -35, -43). They all differ from one another by -8, although the terms are not all multiples of -8; they are offset by 3 (or 5, depending on your perspective). Thus:

$$f(n) = -8n - 3$$

We started with a constant function and then moved to a first-order function. To solve for the general term of the original sequence (4, -7, -26, -53, -88, -131), we must use a second-order relationship. By this point, it may be clear that a sum is needed. A sum has the form:

$$f(n) = \frac{n(n+1)}{2}$$

Modifying the above, we can find:

$$f(n) = -4n(n-1) - 3n + 7$$

fits the terms 4, -7, -26, -53, -88, and -131. This simplifies to:

$$f(n) = -4n^2 + n + 7$$

Using a Best-Fit Model

Sometimes it can be difficult to solve for the general term of a sequence or series by hand. The task can be made much easier if one has access to a graphing calculator or spreadsheet program like Excel or Numbers.

Given the terms 4, -7, -26, -53, -88, and -131 again, these can be converted into points for a Cartesian plot: (1, 4), (2, -7), (3, -26), (4, -53), (5, -88), (6, -131). If these points are graphed, a program can calculate the formula relating all the terms (*Figure 8.2*).

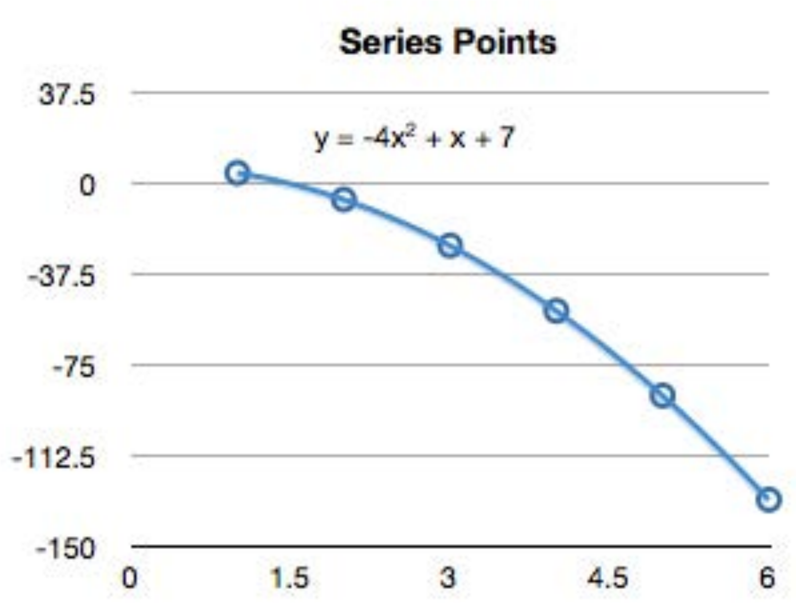


Figure 8.2 Series points fit with trendline

The program Numbers calculated a trendline of y=-4x^2+x+7 to relate points in a series.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/sequences-and-series/finding-the-general-term/

Sums and Series

The summation of all the terms of a sequence is called a series, and many formulae are available for easily calculating large series.

KEY POINTS

- ::A series is merely the sum of the terms of a series. The notation for this operation is to use the capital greek letter sigma, following the general formula: $\sum_{i=m}^{n} x_i = x_m + x_{m+1} + x_{m+2} + \ldots + x_{n-1} + x_n$.
- The sum of an arithmetic series can be calculated using the equation $\frac{n}{2}(t_1 + t_n)$.
- The sum of a geometric series can be calculated using the equation: $S = t_1 \frac{(r^n - 1)}{r - 1}$.

Sums and Series

Summation is the operation of adding a sequence of numbers; the result is their sum or total. If numbers are added sequentially from left to right, any intermediate result is a partial sum, prefix sum, or running total of the summation. The numbers to be summed (called addends, or sometimes summands) may be integers, rational

numbers, real numbers, or complex numbers. Besides numbers, other types of values can be added as well: vectors, matrices, polynomials and, in general, elements of any additive group (or even monoid). For finite sequences of such elements, summation always produces a well-defined sum (possibly by virtue of the convention for empty sums).

A series is merely the sum of the terms of a series. The notation for this operation is to use the capital greek letter sigma, following the general formula:

$$\sum_{i=m}^{n} x_i = x_m + x_{m+1} + x_{m+2} + \ldots + x_{n-1} + x_n$$

where i represents the index of summation; x_i is an indexed variable representing each successive term in the series; m is the lower bound of summation, and n is the upper bound of summation. The "i = m" under the summation symbol means that the index i starts out equal to m. The index, i, is incremented by 1 for each successive term, stopping when i = n.

Arithmetic Series

If you add up all the terms of an **arithmetic** sequence (a sequence in which every entry is the previous entry plus a constant), you have an arithmetic series. For instance:

$10 + 13 + 16 + 19 + 22 + 25 = 105.$

There is a trick that can be used to add up the terms of any arithmetic series. While this trick may not save much time with a 6-item series like the one above, it can be very useful if adding up longer series. The trick is to work from the outside in.

Consider the example given above: $10 + 13 + 16 + 19 + 22 + 25.$

Looking at the first and last terms: $10 + 25 = 35.$

Going in, to the second and next-to-last terms: $13 + 22 = 35.$

Finally, the two inside numbers: $16 + 19 = 35.$

So we can see that the sum of the whole thing is: $3 \cdot 35.$

Pause here and check the following things:

- You understand the calculation that was done for this particular example.
- You understand that this trick will work for any arithmetic series.
- You understand that this trick will not work, in general, for series that are not arithmetic.

If we apply this trick to the generic arithmetic series, we get a formula that can be used to sum up any arithmetic series. Every arithmetic series can be written as follows:

$t_1 + (t_1 + d) + (t_1 + 2d) + \ldots + (t_n - d) + t_n$

If you add the first and last terms, you get $t_1 + t_n$. Ditto for the second and next-to-last terms, and so on. How many such pairs will there be in the whole series? Well, there are n terms, so there are $\frac{n}{2}$ pairs. So the sum for the whole series is $\frac{n}{2}(t_1 + t_n)$.

Geometric Series

If you add up all the terms of a **geometric** sequence (one in which each entry is the previous entry multiplied by a constant), you have a geometric series. The common example of a trend that follows a geometric series is the number of people infected with a virus, as each person passes it to several more, as seen in *Figure 8.3*. The

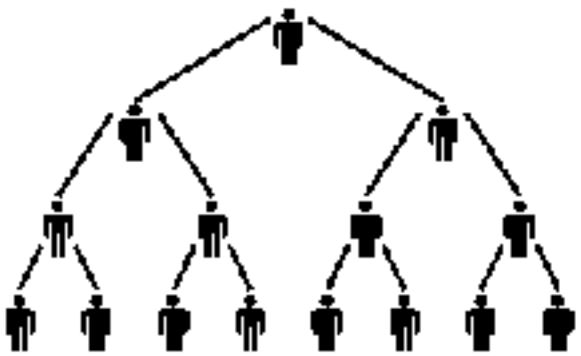


Figure 8.3 Each person infects two more people with the flu virus

So the total number of people infected follows a geometric series.

arithmetic series trick will not work on such a series; however, there is a different trick we can use. As an example, let's find the sum $2 + 6 + 18 + 54 + 162$.

We begin by calling the sum of this series S:

$S = 2 + 6 + 18 + 54 + 162$

Now, if you multiply both sides of this equation by 3 (the same constant that each entry is multiplied by to arrive at the next entry), you get the first equation written below. (The second equation below is just copied from above.)

$3S = 6 + 18 + 54 + 162 + 486$ (*confirm this for yourself!)

$S = 2 + 6 + 18 + 54 + 162$

Here comes the key moment in the trick: subtract the two equations. Notice that the first term of the top equation is the same as the second term of the bottom equation, making much of the subtraction very straightforward. This leaves you with:

$2S = 486 - 2$, so $S = 242$.

Once again, pause to convince yourself that this will work on all geometric series, but only on geometric series.

Finally—once again—we can apply this trick to the generic geometric series to find a formula. So we begin with $t_1 + t_1 r + t_1 r^2 + t_1 r^3 \ldots t_1 r^{n-1}$ and write...

$rS = t_1 r + t_1 r^2 + t_1 r^3 + \ldots + t_1 r^{n-1} + t_1 r^n$ (*confirm this!)

Again, subtracting and soling, we get...

$$rS - S = t_1 r^n - t_1$$

$$S(r-1) = t_1(r^n - 1)$$

$$S = t_1 \frac{(r^n - 1)}{r - 1}$$

So there we have it: a general formula for the sum of any finite geometric series, with the first term t_1, the common ratio r, and a total of n terms.

Notation: Sigma

Sigma notation, which uses the uppercase Greek letter sigma, is used to represent summations—a series of numbers to be added together.

KEY POINTS

- A summation is performed on a series, or list of numbers. Each term is added to the next, resulting in a sum of all terms.
- Sigma notation is used to represent the summation of a series. In this form, the capital Greek letter sigma (Σ) is used. The range of terms in the summation is represented in numbers below and above the Σ; the lowest term is written below and the greatest term is written above.
- Less common forms of sigma notation often leave out specific components of the notation, such as the upper and lower boundaries for the summation to occur—if these boundaries are implicit in the context of the mathematical problem.

Summation is the operation of adding a sequence of numbers, resulting in a sum or total. If numbers are added sequentially from left to right, any intermediate result is a partial sum, prefix sum, or running total of the summation. The numbers to be summed (called addends, or sometimes summands) may be integers, rational numbers, real numbers, or complex numbers.

Aside from numbers, other types of values can be summed: vectors, matrices, polynomials, and generally any elements of an additive group (or even monoid). For finite sequences of such elements, summation always produces a well-defined sum (possibly by virtue of the convention for empty sums).

A summation is performed on a series. A series is a list of numbers—like a sequence—but instead of listing them, you add them all up. For instance, 4+9+3+2+17. (This particular series adds up to 35.)

Sigma Notation

One way to compactly represent a series is with "**sigma** notation," or "summation notation," which looks like this:

$$\sum_{n=3}^{7} n^2$$

The main symbol seen is the uppercase Greek letter for sigma. It indicates a series. To "unpack" this notation, n=3 represents the number at which to start counting (3), and the 7 represents the point at which you stop. For each term, plug that value of n into the given formula (n^2). This particular formula, which we can read as "the sum as n goes from 3 to 7 of n^2" means;

$$3^2 + 4^2 + 5^2 + 6^2 + 7^2$$

More generally, sigma notation can be defined as:

$$\sum_{i=m}^{n} x_i = x_m + x_{m+1} + x_{m+2} + \ldots + x_{n-1} + x_n$$

Where, i represents the index of summation; x_i is an indexed variable representing each successive term in the series; m is the lower bound of summation, and n is the upper bound of summation. The "i = m" under the summation symbol means that the index i starts out equal to m. The index, i, is incremented by 1 for each successive term, stopping when i = n.

Other Forms of Sigma Notation

Informal writing sometimes omits the definition of the index and bounds of summation when these are clear from context, as in:

$$\sum x_i^2 = \sum_{i=1}^{n} x_i^2$$

One often sees generalizations of this notation in which an arbitrary logical condition is supplied, and the sum is intended to be taken over all values satisfying the condition. For example:

$\sum_{0 \le k \le 100} f(k)$ is the sum of f(k) over all (integer) k in the specified range.

$\sum_{x \epsilon S} f(x)$ is the sum of f(x) over all elements x in the set S.

$\sum_{d|n} \mu(d)$ is the sum of μ(d) over all positive integers d dividing n.

Figure 8.4 Summations Involving Exponential Terms

Some summations involving exponential terms

In the summations below *x* is a constant not equal to 1

$$\sum_{i=m}^{n-1} x^i = \frac{x^m - x^n}{1-x} \quad \text{(}m < n\text{; see geometric series)}$$

$$\sum_{i=0}^{n-1} x^i = \frac{1 - x^n}{1-x} \quad \text{(geometric series starting at 1)}$$

$$\sum_{i=0}^{n-1} i x^i = \frac{x - nx^n + (n-1)x^{n+1}}{(1-x)^2}$$

$$\sum_{i=0}^{n-1} i 2^i = 2 + (n-2)2^n \quad \text{(special case when } x = 2\text{)}$$

$$\sum_{i=0}^{n-1} \frac{i}{2^i} = 2 - \frac{n+1}{2^{n-1}} \quad \text{(special case when } x = 1/2\text{)}$$

Formulas relating to summations that you may later encounter in math classes, presented here as a reference.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/sequences-and-series/notation-sigma/

Recursive Definitions

A recursive definition of a function defines its values for some inputs in terms of the values of the same function for other inputs.

KEY POINTS

- In mathematical logic and computer science, a recursive definition, or inductive definition, is used to define an object in terms of itself.
- Most recursive definitions have three foundations: a base case (basis), an inductive clause, and an extremal clause. Often only the inductive clause is discussed, which is the general case, and then as an example one will be asked to demonstrate the inductive clause given a base case or basis.
- The recursive definition for an arithmetic sequence is: $a_n = a_{n-1} + d$. The recursive definition for a geometric sequence is: $a_n = ra_{n-1}$. Again, plugging in an initial value (n_1 =1 for example} will allow the nth term of sequence to be calculated, if r or d is known.

In mathematical logic and computer science, a recursive definition, or inductive definition, is used to define an object in terms of itself. A recursive definition of a function defines values of the functions for some inputs in terms of the values of the same function for other inputs. For example, the factorial function n! is defined by the rules:

$0! = 1$

$(n+1)! = (n+1)n!$

This definition is valid because, for all n, the recursion eventually reaches the base case of 0. Thus the definition is well-founded. The definition may also be thought of as giving a procedure describing how to construct the function n!, starting from n = 0 and proceeding onward with n = 1, n = 2, n = 3 etc.

An inductive definition of a set describes the elements in that set in terms of other elements in another set. For example, one definition of the set N of natural numbers is:

- 1 is in N
- If an element n is in N then n+1 is in N.
- N is the smallest set satisfying these two rules.

There are many sets that satisfy (1) and (2) - for example, the set {1, 1.649, 2, 2.649, 3, 3.649,...} satisfies the definition. However, condition (3) specifies the set of natural numbers by removing the sets with extraneous members.

Form of Recursive Definitions

Most recursive definition have three foundations: a base case (basis), an inductive clause, and an extremal clause.

The difference between a circular definition and a recursive definition is that a recursive definition must always have base cases, cases that satisfy the definition without being defined in terms of the definition itself, and all other cases comprising the definition must be "smaller" (closer to those base cases that terminate the recursion) in some sense. In contrast, a circular definition may have no base case, and define the value of a function in terms of that value itself, rather than on other values of the function. Such a situation would lead to an **infinite regress**.

Examples of Recursive Definitions

All of the elementary functions (addition, multiplication, and exponentiation) can be defined recursively. For addition:

$0 + a = a$

$(1+n) + a = 1 + (n+a)$

Multiplication:

$0a = 0$

$(1+n)a = a + na$

Exponentiation can also be defined recursively:

$a^0 = 1$

$a^{1+n} = aa^n$

Prime Numbers

The set of prime numbers can be defined as the unique set of positive integers satisfying:

- 1 is not a prime number.
- any other positive integer is a prime number if and only if it is not divisible by any prime number smaller than itself.

The primality of the integer 1 is the base case. Checking the primality of any larger integer X by this definition requires knowing the primality of every integer between 1 and X, which is well defined by this definition. That last point can be proved by induction on X, for which it is essential that the second clause says "if and only if". If the clause had said just "if" the primality of, for instance, 4 would not be clear, and the further application of the second clause would be impossible.

Recursive Formulae for Sequences

When discussing arithmetic and quadratic sequences, one may have noticed that the difference between two consecutive terms in the sequence could be written in a general way:

$a_n = a_{n-1} + d$

The above equation is an example of a recursive equation since the nth-term can only be calculated by considering the previous term in the sequence. Compare this with the equation:

$a_n = a_1 + d(n-1)$

One can directly calculate the nth-term of an arithmetic sequence without knowing previous terms.

For quadratic sequences, the difference between consecutive terms is given by equation:

$a_n - a_{n-1} = D(n-2) + d$

Which can be rewritten as:

$a_n = a_{n-1} + D(n-2) + d$

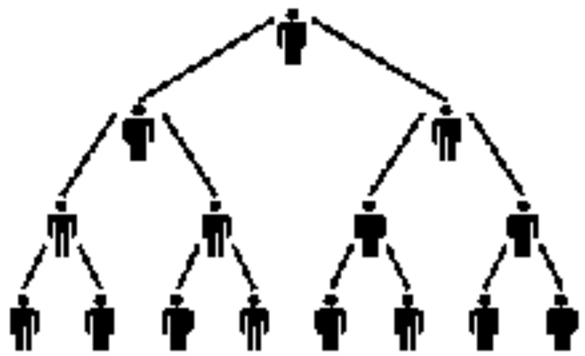
This is then a recursive equation for a quadratic sequence with common second difference, D.

A geometric sequence follows the formula $\frac{a_n}{a_{n-1}} = r$. This can also be stated as a sequence in which every number is equal to the previous number times a constant. For example, the spread of the flu virus often follows a geometric sequence (at least in theory), wherein each person infected will infect two more, such that the terms follow a geometric sequence. Visually, it looks like this: (*Figure 8.5*).

Using this equation, the recursive equation for a geometric sequence is:

$a_n = r \cdot a_{n-1}$

Figure 8.5 The flue virus is a geometric sequence



Each person infects two more people with the flu virus, making the number of recently-infected people the nth term in a geometric sequence.

Recursive equations are extremely powerful: one can work out every term in the series just by knowing previous terms. As can be seen from the examples above, working out and using the previous term a_{n-1} can be a much simpler computation than working out a_n from scratch using a general formula. This means that using a recursive formula when using a computer to work out a sequence would mean the computer would finish its calculations significantly quicker.

Section 2

Arithmetic Sequences and Series

https://www.boundless.com/algebra/sequences-series-and-combinatorics/arithmetic-sequences-and-series/

Arithmetic Sequences

An arithmetic sequence is a sequence of numbers wherein the difference between the consecutive terms is constant.

KEY POINTS

- The behavior of the arithmetic sequence depends on the common difference d.
- Arithmetic sequences can be finite or infinite.
- Finite arithmetic sequences can be summed to make an arithmetic series.

An arithmetic progression, or **arithmetic sequence**, is a sequence of numbers such that the difference between the consecutive terms is constant. For instance, the sequence 5, 7, 9, 11, 13, … is an arithmetic sequence with common difference of 2.

- a_1 The first term of the sequence
- d The common difference of successive terms
- a_n The nth term of the sequence

The behavior of the arithmetic sequence depends on the common difference d.

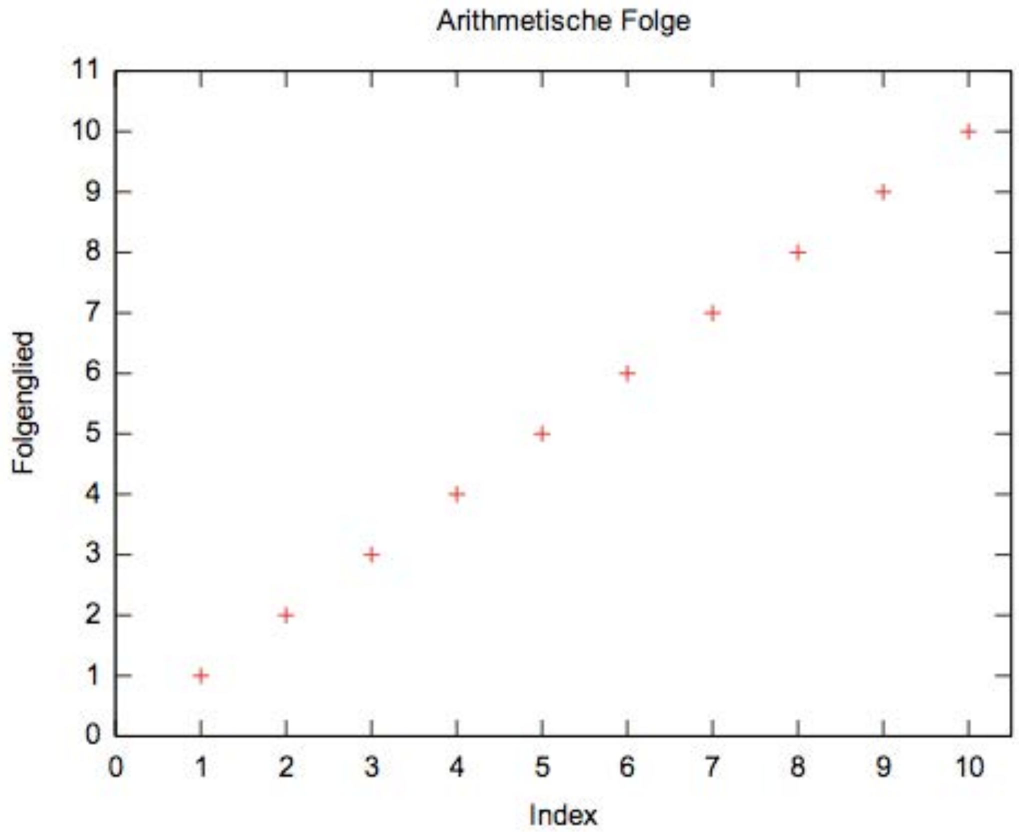
If the common difference, d, is:

- Positive, the sequence will progress towards infinity.
- Negative, the sequence will regress towards negative infinity.

The following equation gives a_n

$$a_n = a_1 + (n - 1) * d$$

Figure 8.6 Arithmetic Sequence



Arithmetic sequences increase each number by a constant amount, similar to a straight line.

Of course, one can always write out each term until getting the answer sought, but if the 50th term is needed doing so can be cumbersome.

An **infinite** arithmetic series is exactly what it sounds like: an infinite series whose terms are in an arithmetic sequence. Examples are $1 + 1 + 1 + 1 + \cdots$ and $1 + 2 + 3 + 4 + \cdots$. The general form for an infinite arithmetic series is:

$$\Sigma(an + b)$$

where n can range from 0 to infinity. If a = b = 0, then the sum of the series is 0. If either a or b is nonzero, then the series has no sum.

Even if one is dealing with an infinite sequence, the sum of that sequence can still be found up to any nth term with the same equation used in a finite arithmetic sequence.

Summing Terms in an Arithmetic Sequence

An arithmetic sequence is a series of numbers in which the difference between consecutive terms is constant.

KEY POINTS

- The sum of the members of a finite arithmetic progression is called an arithmetic series.
- The equation $S_n = (n/2)[2a_1 + (n-1)d]$ can be used to find the sum of any arithmetic sequence up to the nth term.
- Some arithmetic sequences are infinite, and their general form is: $\Sigma(a_n + b)$. While these sequences are infinite, you can still apply the summation equation to find their sum up to a specific nth term.

An arithmetic progression or arithmetic sequence is a sequence of numbers such that the difference between the consecutive terms is constant. For instance, the sequence 5, 7, 9, 11, 13, ... is an arithmetic progression with common difference of 2.

Finite Summation

The sum of the members of a **finite** arithmetic progression is called an arithmetic series.

An arithmetic series can be expressed in two different ways:

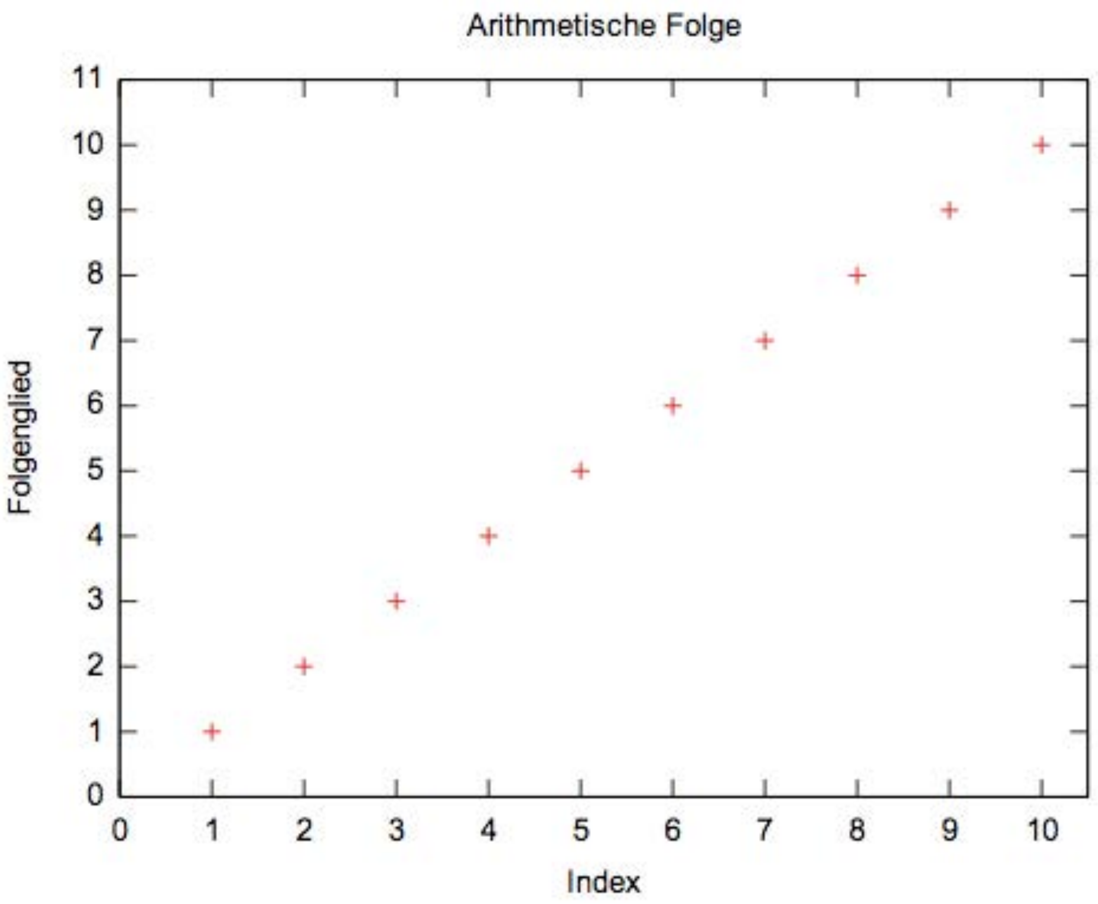
$S_n = a1 + (a1 + d) + (a1 + 2d) + \ldots + (a1 + (n - 2)d) + (a1 + (n - 1)d)$

$S_n = (a_n - (n - 1)d) + (a_n - (n - 2)d) + \ldots + (a_n - 2d) + (a_n - d) + a_n$

Adding both sides of the two equations, all terms involving d cancel:

$2S_n = n(a1 + a_n)$

Figure 8.7 Arithmetic Sequence



Arithmetic sequences increase in value by a set number.

Dividing both sides by 2 produces a common form of the equation:

$S_n = (n/2)(a1 + a_n)$

An alternate form results from reinserting the substitution $a_n = a_1 + (n - 1)d$:

$S_n = (n/2)[2a_1 + (n - 1)d]$

For example, lets say you have an arithmetic progression of 3, 8, 13, 18, 23,..., and you want to know the sum up to the 50th term. Before you can start, you first need to know what information is relevant. In this case you will need:

- The number terms you want to sum (n)
- The first term of the sequence (a_1)
- and the difference in consecutive terms (d)

From the question, we know that n = 50. We can easily see what the first term is, a_1 = 3. And because we know simple math, we can quickly figure out that d = 5. So now, lets apply the equation: $S_{50} = (50/2) * [2(3) + (49)5] = 6275$. That is a lot easier than actually writing out each term and adding them all together.

Infinite Summation

An **infinite** arithmetic series is exactly what it sounds like. It is an infinite series whose terms are in an arithmetic progression. Examples are $1 + 1 + 1 + 1 + \cdots$ and $1 + 2 + 3 + 4 + \cdots$. The general form for an infinite arithmetic series is

$$\sum (a_n + b)$$

where n ranges from 0 to infinity.

If a = b = 0, then the sum of the series is 0. If either a or b is nonzero, then the series diverges, which means it has no sum.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/arithmetic-sequences-and-series/summing-terms-in-an-arithmetic-sequence/

Applications and Problem Solving

Arithmetic series can simplify otherwise complex addition problems by decreasing the number of terms to be added.

KEY POINTS

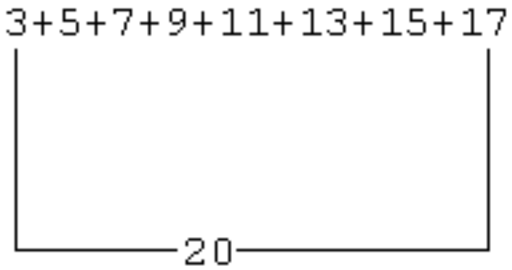
- Arithmetic sequences can sometimes delineate very long lists of numbers that are not practical to add by hand.
- Because the sums of first and last terms, the second and the second-to-last terms, the third and the third-to-last-terms, etc., are all identical, summations of arithmetic sequences can be simplified.
- By simplifying the math for an arithmetic series, one can calculate the sums of large strings of numbers.

Using equations for **arithmetic sequence summation** can greatly facilitate the speed of problem solving. For example, let's say we wanted to write the series of all the even numbers between 50 and 100. This can be written as:

$$\sum_{n=1}^{26} 48 + 2n$$

This can be written in a few different ways, but they all have one thing in common: adding them up by hand would be difficult. There is a quicker solution.

Figure 8.9 Adding a Series



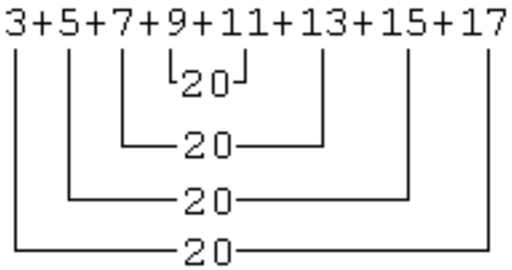
First we should add the first and last terms.

Consider the series: 3+5+7+9+11+13+15+17.

Adding the first term to the last term, 3 + 17 = 20.

Adding the second term to the second-to-last term also amounts to a sum of 20. Adding to ([*Figure 8.9*]), we can see that the third term and third-to-last terms have a similar effect ([*Figure 8.8*]). There are eight terms in 3+5+7+9+11+13+15+17, and they add to four 20's, or 80.

Figure 8.8 Adding a Series Part 2



Adding the first and the last, second term and second to last, etc. all yield the same answer.

This trick applies to all arithmetic series. The reason that the sum of the second pair equaled that of the first pair was that we went up by two on the left, and down by two on the right. As long as you go up by the same amount as you go down, the sum will stay the same—and this is just what happens for arithmetic series.

Remember, that this can be generalized as:

$$S_n = \frac{n}{2} \cdot [2a_1 + (n-1) \cdot d]$$

To apply this to the first summation of all the even numbers between 50 and 100, we would want to add until the 50th term:

$$S_{50} = \frac{50}{2} \cdot [2(50) + (50-1)(2)]$$

$$= 4950$$

Thus, arithmetic series can simplify otherwise complex addition problems!

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/arithmetic-sequences-and-series/applications-and-problem-solving--5/

Boundless is an openly licensed educational resource

Section 3

Geometric Sequences and Series

Geometric Sequences

Summing the First n Terms in an Geometric Sequence

Infinite Geometric Series

Applications and Problem Solving

https://www.boundless.com/algebra/sequences-series-and-combinatorics/geometric-sequences-and-series/

Geometric Sequences

A geometric sequence, is a sequence of numbers where each term after the first is found by multiplying the previous one by the common ratio.

KEY POINTS

- The sum of the terms of a geometric progression, or of an initial segment of a geometric progression, is known as a geometric series.
- The general form of a geometric sequence is: $a, \; ar, \; ar^2, \; ar^3, \; ar^4, \; \ldots$
- The n-th term of a geometric sequence with initial value a and common ratio r is given by: $a^n = ar^{n-1}$.

A geometric progression, also known as a **geometric sequence**, is a sequence of numbers where each term after the first is found by multiplying the previous one by a fixed non-zero number called the common ratio *Figure 8.10*. For example, the sequence 2, 6, 18, 54,... is a geometric progression with common ratio 3. Similarly 10, 5, 2.5, 1.25,... is a geometric sequence with common ratio 1/2. The sum of the terms of a geometric progression, or of an initial segment of a geometric progression, is known as a geometric series.

Figure 8.10 Geometric Sequence

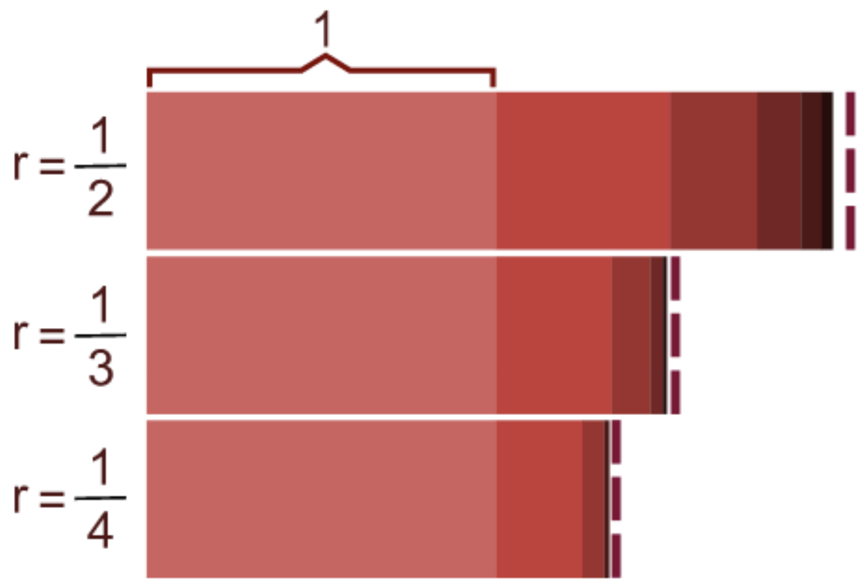


Diagram illustrating three basic geometric sequences of the pattern 1(rn-1) up to 6 iterations deep. The first block is a unit block and the dashed line represents the infinite sum of the sequence, a number that it will forever approach but never touch:,, and, respectively.

Thus, the general form of a geometric sequence is:

$a, \; ar, \; ar^2, \; ar^3, \; ar^4, \; \ldots$

and that of a geometric series is: $a + ar + ar^2 + ar^3 + ar^4 + \ldots$

where r ≠ 0 is the common ratio and a is a scale factor, equal to the sequence's start value.

The n-th term of a geometric sequence with initial value a and common ratio r is given by

$a^n = ar^{n-1}$

Such a geometric sequence also follows the recursive relation

$a^n = ar_{n-1}$ for every integer $n \geq 1$

Generally, to check whether a given sequence is geometric, one simply checks whether successive entries in the sequence all have the same ratio. The common ratio of a geometric series may be negative, resulting in an alternating sequence, with numbers switching from positive to negative and back. For instance: 1, −3, 9, −27, 81, −243,... is a geometric sequence with common ratio −3.

The behavior of a geometric sequence depends on the value of the common ratio.

If the common ratio is:

- Positive, the terms will all be the same sign as the initial term.
- Negative, the terms will alternate between positive and negative.
- Greater than 1, there will be exponential growth towards positive infinity.
- 1, the progression is a constant sequence.
- Between −1 and 1 but not zero, there will be exponential decay towards zero.
- −1, the progression is an alternating sequence (see alternating series)
- Less than −1, for the absolute values there is exponential growth towards positive and negative infinity (due to the alternating sign).

Geometric sequences (with common ratio not equal to −1,1 or 0) show exponential growth or exponential decay, as opposed to the Linear growth (or decline) of an arithmetic progression such as 4, 15, 26, 37, 48, ... (with common difference 11). This result was taken by T.R. Malthus as the mathematical foundation of his Principle of Population. Note that the two kinds of progression are related: exponentiating each term of an arithmetic progression yields a geometric progression, while taking the logarithm of each term in a geometric progression with a positive common ratio yields an arithmetic progression.

An interesting result of the definition of a geometric progression is that for any value of the common ratio, any three consecutive terms a, b and c will satisfy the following equation:

$b^2 = ac$

A geometric series is the sum of the numbers in a geometric progression:

$$\sum_{k=0}^{n} ar^k = \quad ar^0 + ar^1 + ar^2 + ar^3 + \cdots + ar^n$$

We can find a simpler formula for this sum by multiplying both sides of the above equation by 1 – r, and we'll see that

$$(1-r)\sum_{k=0}^{n} ar^k = \quad (1-r)(ar^0 + ar^1 + ar^2 + ar^3 + \cdots + ar^n)$$

$$\begin{aligned} &= ar^0 + ar^1 + ar^2 + ar^3 + \cdots + ar^n \\ &= -ar^1 - ar^2 - ar^3 - \cdots - ar^n - ar^{n+1} \\ &= a - ar^{n+1} \end{aligned}$$

since all the other terms cancel. If r ≠ 1, we can rearrange the above to get the convenient formula for a geometric series:

$$\sum_{k=0}^{n} ar^k = \quad \frac{a(1-r^{n+1})}{1-r}$$

Summing the First n Terms in an Geometric Sequence

By utilizing the common ratio and the first term of the sequence, we can sum the first n terms: $s = a\frac{1-r^n}{1-r}$.

KEY POINTS

- The terms of a geometric series form a geometric progression, meaning that the ratio of successive terms in the series is constant.
- The behavior of the terms depends on the common ratio r.
- The sum of a geometric series is finite as long as the terms approach zero; as the numbers near zero, they become insignificantly small, allowing a sum to be calculated despite the series being infinite. The sum can be computed using the self-similarity of the series.

Geometric series are one of the simplest examples of infinite series with finite sums, although not all of them have this property. Historically, geometric series played an important role in the early development of calculus, and they continue to be central in the study of convergence of series. Geometric series are used throughout mathematics, and they have important applications in

Common ratio	Example
10	4 + 40 + 400 + 4000 + 40,000 + ···
1/3	9 + 3 + 1 + 1/3 + 1/9 + ···
1/10	7 + 0.7 + 0.07 + 0.007 + 0.0007 + ···
1	3 + 3 + 3 + 3 + 3 + ···
−1/2	1 − 1/2 + 1/4 − 1/8 + 1/16 − 1/32 + ···
−1	3 − 3 + 3 − 3 + 3 − ···

Figure 8.11
Geometric Series

Geometric series with different common ratios

physics, engineering, biology, economics, computer science, queueing theory, and finance.

The terms of a geometric series form a geometric progression, meaning that the ratio of successive terms in the series is constant.

The following table shows several geometric series with different common ratios (*Figure 8.11*). The behavior of the terms depends on the common ratio r:

- If r is between −1 and +1, the terms of the series become smaller and smaller, approaching zero in the limit, and the series converges to a sum. In the case above, where r is one half, the series has a sum of one.
- If r is greater than one or less than minus one, the terms of the series become larger and larger in magnitude. The sum of the terms also gets larger and larger, and the series has no sum. (The series diverges.)
- If r is equal to one, all of the terms of the series are the same. The series diverges. If r is minus one, the terms take two values alternately (e.g., 2, −2, 2, −2, 2,...). The sum of the terms oscillates between two values (e.g. 2, 0, 2, 0, 2,...). This is a different type of divergence and again the series has no sum.

The sum of a geometric series is finite as long as the terms approach zero; as the numbers near zero, they become insignificantly small, allowing a sum to be calculated despite the series being infinite. The sum can be computed using the self-similarity of the series.

For example:

$$s = 1 + \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \ldots$$

This series has common ratio 2/3. If we multiply through by this common ratio, then the initial 1 becomes a 2/3, the 2/3 becomes a 4/9, and so on:

$$\frac{2}{3}s = \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \frac{16}{81} + \ldots$$

This new series is the same as the original, except that the first term is missing. Subtracting the new series (2/3)s from the original series, s cancels every term in the original but the first:

$s - \frac{2}{3}s = 1$, so $s = 3$

A similar technique can be used to evaluate any self-similar expression.

For $r \neq 1$, the sum of the first n terms of a geometric series is:

$$a + ar + ar^2 + ar^3 + \cdots + ar^{n-1} = \sum_{k=0}^{n-1} ar^k = a\frac{1-r^n}{1-r},$$

Where a is the first term of the series, and r is the common ratio. We can derive this formula as follows:

Let $s = a + ar + ar^2 + ar^3 + \cdots + ar^{n-1}$

Then $rs = ar + ar^2 + ar^3 + ar^4 \cdots + ar^n$

Then $s - rs = a - ar^n$

Then $s(1-r) = a(1-r^n)$

So $s = a\frac{1-r^n}{1-r}$

Therefore, by utilizing the common ratio and the first term of the sequence, we can sum the first n terms of a sequence.

Infinite Geometric Series

Geometric series are one of the simplest examples of infinite series with finite sums.

KEY POINTS

- The sum of a geometric series is finite as long as the terms approach zero; as the numbers near zero, they become insignificantly small, allowing a sum to be calculated despite the series being infinite.
- Historically, geometric series played an important role in the early development of calculus, and they continue to be central in the study of convergence of series.
- The general form of an infinite geometric series is: $\sum_{n=0}^{\infty} z^n$.

Geometric series, or infinite series whose terms are in a geometric progression, are one of the simplest examples of infinite series with finite sums, although not all of them have this property. Historically, geometric series played an important and indispensable role in the early development of calculus, and even today they continue to be central in the study of convergence of series. Geometric series are used throughout mathematics, but they also have important applications in physics, engineering, biology, economics, computer science, queueing theory, and finance.

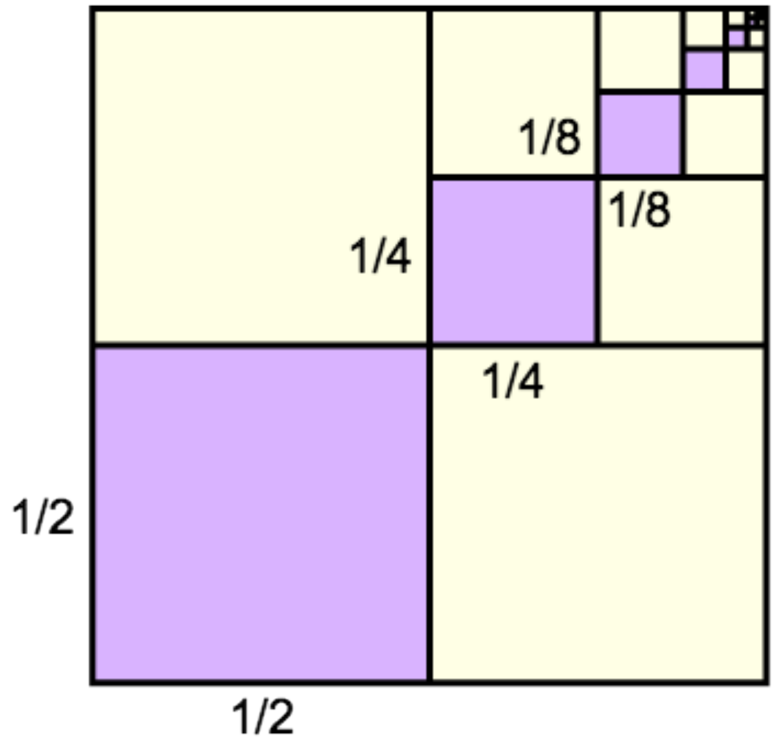


Figure 8.12 Geometric Series

Each of the purple squares is obtained by multiplying the area of the next larger square by 1/4 (1/2×1/2=1/4, 1/4×1/4=1/16). The sum of the areas of the purple squares is one third of the area of the large square.

A geometric series is a series with a constant ratio between successive terms, as seen in the *Figure 8.12*. For example, the following series: $\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots=\sum_{n=0}^{\infty}\frac{1}{2^n}$

is geometric, because each successive term can be obtained by multiplying the previous term by 1 / 2. The general form of a geometric series is $\sum_{n=0}^{\infty} z^n$.

The sum of a geometric series is finite as long as the terms approach zero; as the numbers near zero, they become insignificantly small,

allowing a sum to be calculated despite the series being infinite. The sum can be computed using the self-similarity of the series.

An infinite geometric series is an infinite series whose successive terms have a common ratio. Such a series converges if and only if the absolute value of the common ratio is less than one ($|r| < 1$).

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/geometric-sequences-and-series/infinite-geometric-series/

Applications and Problem Solving

Geometric series have applications in math and science, and are one of the simplest examples of infinite series with finite sums.

KEY POINTS

- A repeating decimal can be viewed as a geometric series whose common ratio is a power of 1/10.
- Archimedes used the sum of a geometric series to compute the area enclosed by a parabola and a straight line.
- The interior of the Koch snowflake is a union of infinitely many triangles. In the study of fractals, geometric series often arise as the perimeter, area or volume of a self-similar figure.

Geometric series played an important role in the early development of calculus, and continue as a central part of the study of convergence of series. Geometric series are used throughout mathematics; they have important applications in physics, engineering, biology, economics, computer science, queueing theory and finance.

Geometric series are one of the simplest examples of infinite series with finite sums, although not all of them have this property.

Repeating Decimal

A repeating decimal can be thought of as a geometric series whose common ratio is a power of 1/10. For example:

$$0.7777\ldots = \frac{7}{10} + \frac{7}{100} + \frac{7}{1000} + \frac{7}{10000} + \cdots$$

The formula for the sum of a geometric series can be used to convert the decimal to a fraction:

$$0.7777\ldots = \frac{a}{1-r} = \frac{7/10}{1-1/10} = \frac{7}{9}$$

The formula works for any repeating term. For example:

$$0.123412341234\ldots = \frac{a}{1-r} = \frac{1234/1000}{1-1/1000} = \frac{1234}{9999}$$

$$0.0909090909 = \frac{.09}{99} = \frac{1}{11}$$

$$0.143814381438 = \frac{.1438}{9999}$$

$$0.9999 = \frac{9}{9} = 1$$

That is, a repeating decimal with repeat length n is equal to the quotient of the repeating part (as an integer) and 10^n - 1.

Archimedes' Quadrature of the Parabola

Archimedes used the sum of a geometric series to compute the area enclosed by a parabola and a straight line (*Figure 8.13*). His method was to dissect the area into an infinite number of triangles. Archimedes' Theorem states that the total area under the parabola is 4/3 of the area of the blue triangle. He determined that each green triangle has 1/8 the area of the blue triangle, each yellow triangle has 1/8 the area of a green triangle, and so forth. Assuming that the blue triangle has area 1, the total area is an infinite series:

$$1 + 2\left(\frac{1}{8}\right) + 4\left(\frac{1}{8}\right)^2 + 8\left(\frac{1}{8}\right)^3 + \cdots$$

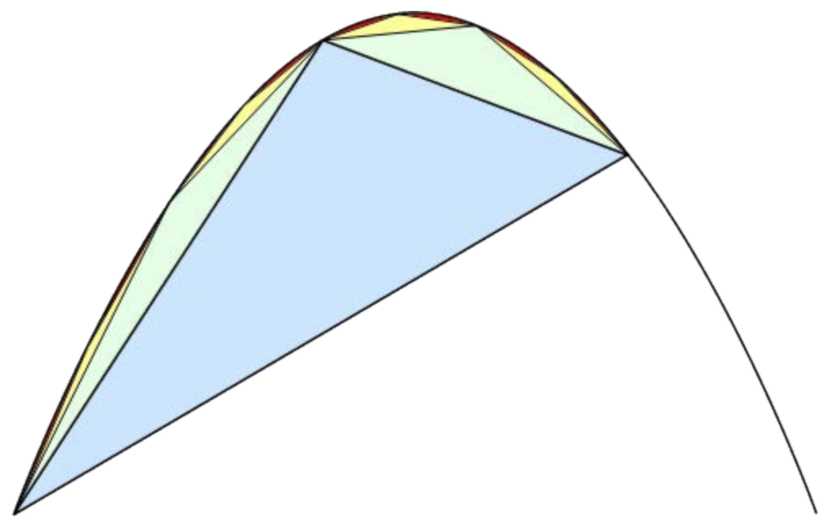


Figure 8.13 Archimedes Theorem

Archimedes' dissection of a parabolic segment into infinitely many triangles.

The first term represents the area of the blue triangle, the second term the areas of the two green triangles, the third term the areas of the four yellow triangles, and so on. Simplifying the fractions gives:

$$1 + \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots$$

This is a geometric series with common ratio 1/4 and the fractional part is equal to 1/3.

Fractal Geometry

The interior of the Koch snowflake is comprised of an infinite amount of triangles, as shown in *Figure 8.14*. In the study of **fractals**, geometric series often arise as the perimeter, area, or volume of a self-similar figure. For example, the area inside the Koch snowflake can be described as the union of many infinitely equilateral triangles. Each side of the green triangle is exactly 1/3 the size of a side of the large blue triangle, and therefore has exactly 1/9 the area. Similarly, each yellow triangle has 1/9 the area of a green triangle, and so forth. Taking the blue triangle as a unit of area, the total area of the snowflake is:

$$1 + 3\left(\frac{1}{9}\right) + 12\left(\frac{1}{9}\right)^2 + 48\left(\frac{1}{9}\right)^3 + \cdots$$

The first term of this series represents the area of the blue triangle, the second term the total area of the three green triangles, the third term the total area of the twelve yellow triangles, and so forth. Excluding the initial 1, this series is geometric with constant ratio r = 4/9. The first term of the geometric series is a = 3(1/9) = 1/3, so the sum is

$$1 + \frac{a}{1-r} = 1 + \frac{\frac{1}{3}}{1 - \frac{4}{9}} = \frac{8}{5}$$

Thus the Koch snowflake has 8/5 of the area of the base triangle.

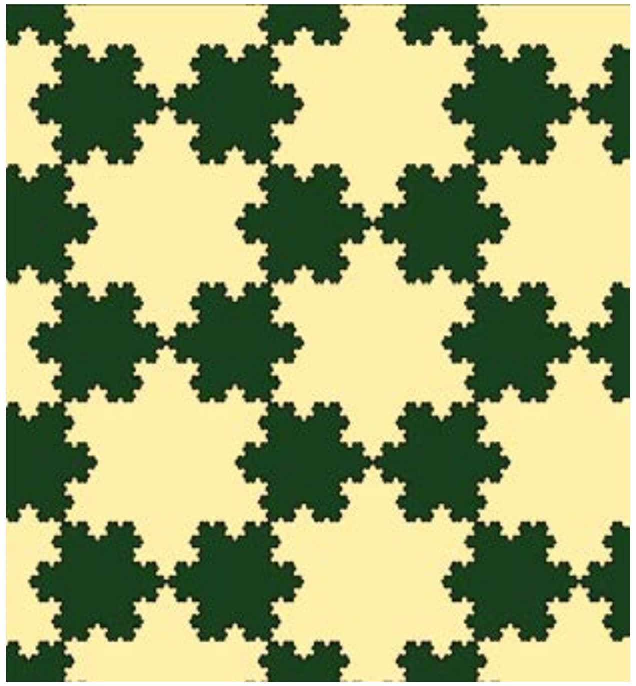


Figure 8.14 Koch Snowflake

The interior of a Koch snowflake is comprised of an infinite amount of triangles.

Zeno's Paradoxes

Zeno's Paradoxes reveal that the convergence of a geometric series means that a sum involving an infinite number of terms can be finite. Simply stated, Zeno's paradox says: There is a point, A, that wants to move to another point, B. If A only moves half of the distance between it and point B at a time, it will never get there, because you can continue to divide the remaining space in half for ever. However, we know this is not true, otherwise we would never reach our destinations.

Economics

In economics, geometric series are used to represent the present value of an **annuity** (a sum of money to be paid in regular intervals).

Geometric Power Series

Assume this formula for a geometric series:

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots$$

This can be interpreted as a power series in the Taylor's theorem sense, converging where $|x| < 1$. One can thus extrapolate to obtain other power series.

Section 4

Mathematical Inductions

https://www.boundless.com/algebra/sequences-series-and-combinatorics/mathematical-inductions/

Sequences of Statements

Sequences of statements are logical, ordered groups of statements that are important for mathematical induction.

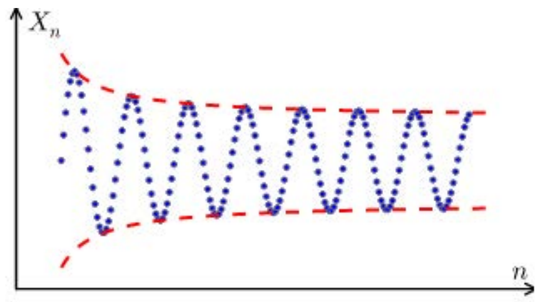
KEY POINTS

- A sequence is an ordered list of objects or events. Like a set, it contains members, but unlike a set, the order of the members matters.
- A sequence of statements refers to the progression of logical implications of one statement.
- Sequences of statements are important for mathematical inductions, which rely on infinite sequences of statements.

In mathematics, a sequence is an ordered list of objects (or events). Like a **set**, it contains members (also called "elements," or "terms"). The number of ordered elements (possibly infinite) is called the "length of the sequence." Unlike a set, order matters in sequences and exactly the same elements can appear multiple times at different positions in the sequence. A sequence is a discrete function.

For example, (M, A, R, Y) is a sequence of letters that differs from (A, R, M, Y), although the composition is the same; the ordering differs. Also (1, 1, 2, 3, 5, 8) is a valid sequence despite including a repeating term (1). Sequences can be finite, as in this example, or infinite, such as the sequence of all even positive integers (2, 4, 6,...). Finite sequences are sometimes known as strings or words and infinite sequences as streams. The empty sequence () is included in most notions of sequence, but may be excluded depending on the context.

Figure 8.15 Sequence



Part of an infinite sequence of real numbers (in blue). This sequence is neither increasing, nor decreasing, nor convergent. It is, however, bounded within the two dashed lines.

A sequence of statements refers to the progression of logical implications of one statement. In algebra, a "statement" usually refers to an equation that contains an equal sign. Sequences of statements are necessary for mathematical induction. Mathematical induction is a method of mathematical proof typically used to establish that a given statement is true for all natural numbers (positive integers). It is done by proving that the first statement in the infinite sequence of statements is true, and then proving that if any one statement in the infinite sequence of statements is true, then so is the next one.

Proving Infinite Sequences of Statements

Proving an infinite sequence of statements is necessary for proof by induction, a rigorous form of deductive reasoning.

KEY POINTS

- Mathematical induction is a method of mathematical proof typically used to establish that a given statement is true for all natural numbers (positive integers).
- It is done by proving that the first statement in the infinite sequence of statements is true, and then proving that if any one statement in the infinite sequence of statements is true, then so is the next one.
- Proving an infinite sequences of statements can be understood in the context of the domino effect, which by nature mediates a sequential and predictable order of events.

Mathematical induction is a method of mathematical proof typically used to establish that a given statement is true for all natural numbers (positive integers). It is done by proving that the first statement in the infinite sequence of statements is true, and then proving that if any one statement in the infinite sequence of statements is true, then so is the next one.

Figure 8.16
Dominoes
Mathematical induction can be informally illustrated by reference to the sequential effect of falling dominoes.

Mathematical induction should not be misconstrued as a form of **inductive reasoning**, which is considered non-rigorous in mathematics (see problem of induction for more information). In fact, mathematical induction is a form of rigorous deductive reasoning.

The simplest and most common form of mathematical induction proves that a statement involving a natural number n holds for all values of n. The proof consists of two steps:

1. The basis (base case): showing that the statement holds when n is equal to the lowest value that n is given in the question. Usually, n = 0 or n = 1.
2. The inductive step: showing that if the statement holds for some n, then the statement also holds when n + 1 is substituted for n.

The assumption in the inductive step that the statement holds for some n, is called the induction hypothesis (or inductive hypothesis). To perform the inductive step, one assumes the induction hypothesis and then uses this assumption to prove the statement for n + 1.

The choice between n = 0 and n = 1 in the base case is specific to the context of the proof: If 0 is considered a natural number, as is common in the fields of combinatorics and mathematical logic, then n = 0. If, on the other hand, 1 is taken as the first natural number, then the base case is given by n = 1.

This method works by first proving the statement is true for a starting value, and then proving that the process used to go from one value to the next is valid. If these are both proven, then any value can be obtained by performing the process repeatedly. It may be helpful to think of the domino effect. If one is presented with a long row of dominoes standing on end, one can be sure that:

1. The first domino will fall
2. Whenever a domino falls, its next neighbor will also fall

So it is concluded that all of the dominoes will fall, and that this fact is inevitable.

Section 5

Combinatorics

https://www.boundless.com/algebra/sequences-series-and-combinatorics/combinatorics/

Permutations

A permutation of a set of objects is an arrangement of those objects into a particular order; the numbers of permutations can be counted.

KEY POINTS

- Informally, a permutation of a set of objects is an arrangement of those objects into a particular order. For example, there are six permutations of the set {1,2,3}, namely (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), and (3,2,1), or with colored balls.
- The number of permutations of n distinct objects is $n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1$, which number is called "n factorial" and written "n!".
- When deciding permutations of a subset from a larger set, it is often useful to divide one factorial by another to determine the number of permutations possible. For example, the first six cards from a deck of cards would have $\frac{52!}{46!}$ permutations possible, or about 14.7 billion.

In the game of "Solitaire", also known as "Patience" or "Klondike", seven cards are dealt out at the beginning, as shown in *Figure 8.17*: one face-up, and the other six face-down. A complete card deck has 52 cards. Assuming that the only card that is seen is the 7 of spades, how many possible "hands" (the other six cards) could be showing underneath? What makes this a "**permutations**" problem is that the order matters: if an ace is hiding somewhere in those six cards, it makes a difference whether the ace is on the first position, the second, etc. Permutations problems can always be addressed as an example of the multiplication rule, with one small twist.

Figure 8.17 One stack of cards in a game of solitaire

To find out how many possible combinations of cards there are below the seven, we use the concept of permutations to calculate the possible arrangements of cards.

Question: How many cards might be in the first position, directly under the showing 7?

Answer: 51. That card could be anything except the 7 of spades.

Question: For any given card in first position, how many cards might be in second position?

Answer: 50. The seven of spades, and the next card, are both "spoken for." So there are 50 possibilities left in this position.

Question: So how many possibilities are there for the first two positions combined?

Answer: 51×50, or 2,550.

Question: So how many possibilities are there for all six positions?

Answer: 51×50×49×48×47×46, or approximately 1.3×10^{10}; about 10 billion possibilities!

This result can be expressed, and typed into a calculator, more concisely by using **factorials**.

A “factorial”, written with an exclamation mark, means “multiply all the numbers from 1 up to this number.” So 5! means 1×2×3×4×5=120.

$\frac{7!}{5!}$ can also be written as, $\frac{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7}{1\cdot2\cdot3\cdot4\cdot5}$. Most of the terms cancel, leaving only 6×7=42.

In another example, $\frac{51!}{45!}$, if all of the terms are written out, the first 45 terms cancel, leaving 46×47×48×49×50×51, which is the number of permutations wanted. Therefore, instead of typing into a calculator six numbers to multiply, or sixty numbers or six hundred depending on the problem, the answer to a permutation problem can be found by dividing two factorials. In many calculators, the factorial option is located under the “probability” menu for this reason.

General Considerations

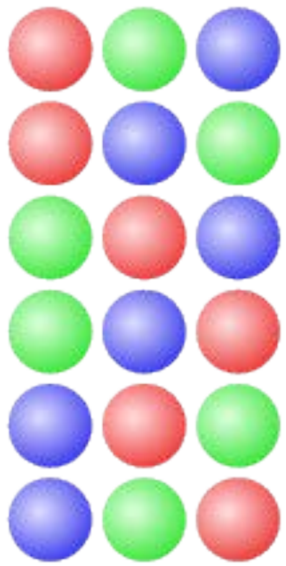


Figure 8.18 The 6 permutation of 3 balls

If one has three different colored balls, there are six distinct ways to organize them into a different order, as shown.

In mathematics, the notion of permutation is used with several slightly different meanings, all related to the act of permuting (rearranging) objects or values. Informally, a permutation of a set of objects is an arrangement of those objects into a particular order. For example, there are six permutations of the set {1,2,3}, namely (1,2,3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), and (3,2,1), or with colored balls, as in *Figure 8.18*. One might define an anagram of a word as a permutation of its letters. The study of permutations in this sense generally belongs to the field of combinatorics.

The number of permutations of n distinct objects is $n\times(n-1)\times(n-2)\times\cdots\times2\times1$, which number is called "n factorial" and written "n!".

Permutations occur, in more or less prominent ways, in almost every domain of mathematics. They often arise when different orderings on certain finite sets are considered, possibly only because one wants to ignore such orderings and needs to know how many configurations are thus identified. For similar reasons

permutations arise in the study of sorting algorithms in computer science.

Permutations: Notation; n Objects Taken k at a Time

A permutation is an arrangement of objects in a specific order; it is one of many possible ways to permute the set of objects.

KEY POINTS

- If all objects in consideration are distinct, they can be arranged in n! permutations, where n represents the number of objects.
- If not all the objects in a set of n unique elements are chosen, the above formula can be modified to: where k represents the number of selected elements.
- When solving for quotients of factorials, the terms of the denominator can cancel with the terms of the numerator, thus eliminating perhaps the majority of terms to be multiplied.

To permute objects is to rearrange them. A **permutation** is an arrangement of objects in a specific order; it is one of many possible ways to permute the set of objects.

For example, there are six possible permutations of the numbers 1, 2, and 3. These are: (1, 2, 3), (1,3,2), (2,1,3), (2,3,1), (3,1,2), and (3,2,1).

If all objects in consideration are distinct, they can be arranged in n! permutations, where n represents the number of objects. The quantity n! is to be read "n **factorial**," meaning it equals the product:

$$n(n-1) \cdot (n-2) \cdot (n-3) \cdot \ldots 2 \cdot 1$$

If not all the objects in a set of n unique elements are chosen, the above formula can be modified for a selection of k elements:

$$\frac{n!}{(n-k)!}$$

Knowing the formula for calculating permutations can help solve quandaries that would be otherwise near-impossible to determine. While it's easy enough to count the six possible permutations of three differently-colored balls, *Figure 8.18* counting the possible seven-card hands in a game of cards is an altogether different level of difficulty. Plugging in 52 (the number of unique cards in a deck) for n and k (the size of the hand) for k, we can find:

$$\frac{52!}{(52-7)!}$$

$$\frac{52!}{48!}$$

6497400

Thus, there are 6497400 possible seven-card hands that can be drawn from a deck of cards.

The above can be solved using a calculator, or by hand. Because 52! and 48! both contain terms 48*47*46*45*44...etc, those terms will cancel. Ultimately, the equation can be simplified to 52*51*50*49.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/combinatorics/permutations-notation-n-objects-taken-k-at-a-time/

Permutations of Nondistinguishable Objects

The expression revealing number of permutations of distinct items can be modified if not all items in a set are distinct.

KEY POINTS

- Some sets include repetitions of certain elements. In these cases, the number of possible permutations of the items cannot be expressed by n!, where n represents the number of elements, because this calculation would include a multiplicity of possible states.
- To correct for the "multiplicity" of certain permutations, divide the factorial of the total number of elements by the product of the factorials of the number of each repeated element.
- The expression for number of permutations with repeated elements is: $\frac{n!}{n_1!n_2!n_3!\dots}$ where n is the total number of terms in a sequence n_1, n_2, and n_3 are the number of repetitions of different elements.

The number of possible **permutations** of a set of n distinct elements n!, or n(n-1)(n-2)...(2)(1).

This can be easily tested. The number 1 can be arranged in just one (1!) way. The numbers 1 and 2 can be arranged in two (2!) ways: (1, 2) and (2, 1). The numbers 1, 2, and 3 can be arranged in six (3!) ways: (1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), and (3, 2, 1). This rule holds true for sets of any size, so long as the elements are all distinct. But what if some elements are repeated?

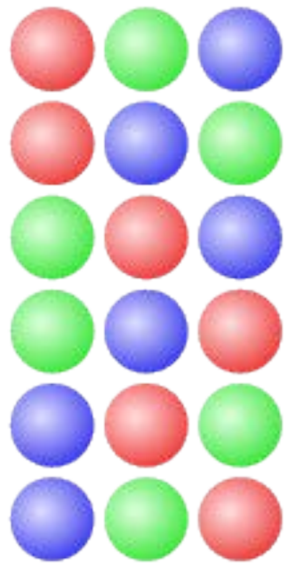


Figure 8.19 The 6 permutation of 3 balls

Imagine replacing each red ball with a green ball. In such a case, three of the six permutations would be repeated. Thus, there would only be three true, distinct permutations.

Repetition of some elements complicates the calculation of permutations, because it allows for there to be multiple ways in which a specific order of elements can be arranged. For example, given the numbers 1, 3, and 3 in a set, there are two ways to obtain the order (3, 1, 3).

To correct for the "**multiplicity**" of certain permutations, we must divide the factorial of the total number of elements by the product of the factorials of the number of each repeated element. This can generally be represented as:

$$\frac{n!}{n_1!n_2!n_3!\dots}$$

Where n is the total number of terms in a sequence n_1, n_2, and n_3 are the number of repetitions of different elements.

Consider the set of numbers: (15, 17, 24, 24, 28)

There are five terms, so n=5. However, two (the number 24) are the same. Thus, the number of possible distinct permutations in the set is:

$$\frac{5!}{2!} = 60$$

The same logic can apply to more complicated systems. Consider the set: (0, 0, 0, 2, 4, 4, 7, 7, 7, 7, 7, 8, 8).

In total, there are 13 elements. These include many repetitions: 0 is seen three times, 4 and 8 each are observed twice, and there are five instances of the number 7. Thus, the number of possible distinct permutations can be calculated by:

$$\frac{13!}{2!\cdot 2!\cdot 3!\cdot 5!} = 2162160$$

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/combinatorics/permutations-of-nondistinguishable-objects/

Combinations

A combination is a way of selecting several things out of a larger group, where (unlike permutations) order does not matter.

KEY POINTS

- A combination is a mathematical concept where one counts the number of ways one can select several elements out of a larger group.
- Unlike a permutation, when determining the number of combinations, order does not matter.
- Formally, a k-combination of a set S is a subset of k distinct elements of S. If the set has n elements the number of k-combinations is equal to the binomial coefficient: $\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots 1}$, which can be written using factorials as $\frac{n!}{k!(n-k)!}$ whenever $k \leq n$ and which is zero when k>n.

Using an example with a deck of 52 cards, deal out a poker hand (5 cards). How many possible poker hands are there?

At first glance, this seems like a minor variation on the Solitaire question (how many ways can be make a stack of six cards, a permutation question). The only real difference is that there are five cards instead of six. But in face, there is a more important difference: order does not matter. One does not want to count "Ace-King-Queen-Jack-Ten of spades" and "Ten-Jack-Queen-King-Ace of spades" separately. They are the same poker hand.

To approach such question, begin with the permutations question: how many possible poker hands are there, if order does matter? 52×51×50×49×48, or $\frac{52!}{47!}$. One has to count every possible hand many different times in this calculation. How many times?

The key insight is that this second question—"How many different times is one counting, for instance, Ace-King-Queen-Jack-Ten of spades?"—is itself a permutations question. It is the same as the question "How many different ways can these five cards be rearranged in a hand?" There are five possibilities for the first card, or each of these, four for the second, and so on. The answer is 5, which is 120. So, since every possible hand has been counted 120 times, divide our earlier result by 120 to find that there are $\frac{52!}{(47!)(5!)}$, or about 2.6 Million possibilities.

The question—"how many different 5-card hands can be made from 52 cards?"—turns out to have a surprisingly large number of applications. Consider the following questions:

- A school offers 50 classes. Each student must choose 6 of them to fill out a schedule. How many possible schedules can be made?
- A basketball team has 12 players, but only 5 will start. How many possible starting teams can they field?
- Your computer contains 300 videos, but you can only fit 10 of them on your iPod. How many possible ways can you load your iPod?

Each of these is a combinations question, and can be answered exactly like the card scenario. Since this type of question comes up in so many different contexts, it is given a special name and symbol. The last question would be referred to as "300 choose 10" and written $\binom{300}{10}$. It is calculated as $\frac{300!}{(290!)(10!)}$, for reasons explained above.

General considerations

In mathematics, a combination is a way of selecting several things out of a larger group, where (unlike permutations) order does not matter. In smaller cases, it is possible to count the number of combinations. For example given three fruit, say an apple, orange and pear, there are three combinations of two that can be drawn from this set: an apple and a pear; an apple and an orange; or a pear and an orange. More formally a k-combination of a set S is a subset of k distinct elements of S. If the set has n elements the number of k-combinations is equal to the **binomial coefficient** (*Figure 8. 20*).

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k(k-1)\dots1}$$

Which can be written using factorials as $\frac{n!}{k!(n-k)!}$ whenever $k \leq n$, and which is zero when $k > n$. The set of all k-combinations of a set S is sometimes denoted by $\binom{S}{k}$.

Combinations can refer to the combination of n things taken k at a time without or with repetitions. In the above example, repetitions were not allowed. If, however, it was possible to have two of any one kind of fruit there would be 3 more combinations: one with two apples, one with two oranges, and one with two pears.

Number of k-combinations

The number of k-combinations from a given set S of n elements is often denoted in elementary combinatorics texts by C(n, k), or by a variation such as C_k^n, ${}_nC_k$, ${}^nC_k$, or even C_n^k (the latter form is standard in French, Russian, and Polish texts). The same number, however, occurs in many other mathematical contexts, where it is denoted by

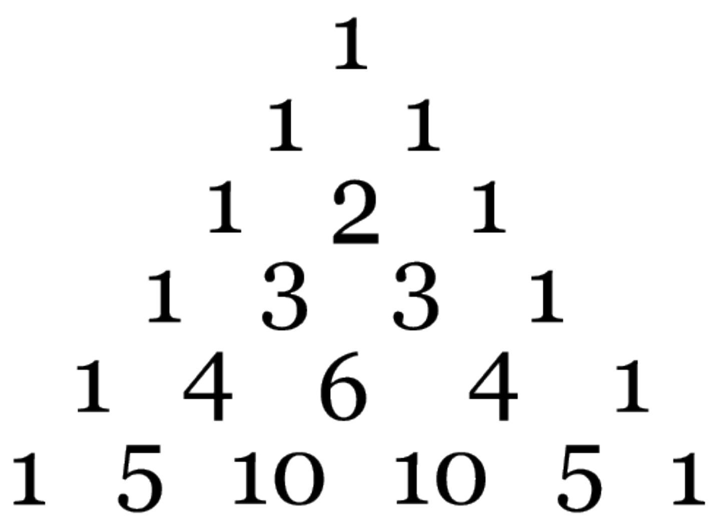


Figure 8.20
Pascal's triangle
The binomial coefficient can be arranged to form Pascal's trangle. The first row is the terms for (x+y)^0, the second for (x+y)^1, the third for (x=y)^2, which is 1x^2+2xy+y^2.

$\binom{n}{k}$ (often read as "n choose k"). Notably, it occurs as coefficient in the binomial formula, hence its name, binomial coefficient. One can define $\binom{n}{k}$ for all natural numbers k at once by the relation

$$(1+X)^n = \Sigma_{k\geq 0}\binom{n}{k}X^k$$

From which it is clear that $\binom{n}{0} = \binom{n}{n} = 1$ and $\binom{n}{k} = 0$ for $k > n$.

Section 6

The Binomial Theorem

https://www.boundless.com/algebra/sequences-series-and-combinatorics/the-binomial-theorem/

Binomial Expansions and Pascal's Triangle

The binomial theorem, which uses Pascal's triangles to determine coefficients, describes the algebraic expansion of powers of a binomial.

KEY POINTS

- According to the theorem, it is possible to expand the power (x + y)n into a sum involving terms of the form axbyc, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a of each term is a specific positive integer depending on n and b.
- Using summation notation, the binomial theorem can be expressed as: $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.
- The rows of Pascal's triangle are conventionally enumerated starting with row n = 0 at the top. The entries in each row are numbered from the left beginning with k = 0 and are usually staggered relative to the numbers in the adjacent rows.

The **binomial** theorem describes the algebraic expansion of powers of a binomial. Essentially, it demonstrates what happens when you multiply a binomial by itself (as many times as you want). According to the theorem, it is possible to expand the power $(x+y)^n$ into a sum involving terms of the form ax^by^c, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a of each term is a specific positive integer depending on n and b. When an exponent is zero, the corresponding power is usually omitted from the term. For example,

$(x+y)^4 = x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4$

The coefficient a in the term of x^by^c is known as the binomial coefficient or (the two have the same value). These coefficients for varying n and b can be arranged to form Pascal's triangle. These numbers also arise in combinatorics, where gives the number of different combinations of b elements that can be chosen from an n-element set.

According to the theorem, it is possible to expand any power of x + y into a sum of the form:

$$(x+y)^n = \binom{n}{0} x^n y^0 + \binom{n}{1} x^{n-1} y^1 + \binom{n}{2} x^{n-2} y^2 + \ldots + \binom{n}{n-1} x^1 y^{n-1} + \binom{n}{n} x^0 y^n$$

where each is a specific positive integer known as binomial coefficient. This formula is also referred to as the Binomial Formula or the Binomial Identity. Using summation notation, it can be written as:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

Pascal's Triangle

Pascal's triangle, *Figure 8.21*, determines the coefficients which arise in binomial expansions. For an example, consider the expansion:

$$(x+y)^2 = x^2 + 2xy + y^2 = 1x^2y^0 + 2x^1y^1 + 1x^0y^2$$

Notice the coefficients are the numbers in row two of Pascal's triangle: 1, 2, 1. In general, when a binomial like x + y is raised to a positive integer power we have:

$$(x+y)^n = a_0x^n + a_1x^{n-1}y + a_2x^{n-2}y^2 + \cdots a_{n-1}xy^{n-1} + a_ny^n$$

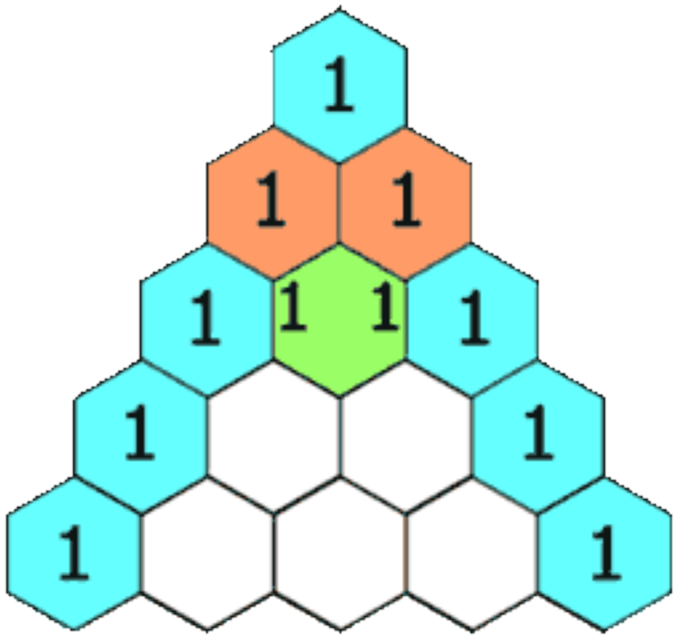


Figure 8.21
Pascal's Triangle
Each number in the triangle is the sum of the two directly above it.

where the coefficients ai in this expansion are precisely the numbers on row n of Pascal's triangle. In other words,

$$a_i = \binom{n}{i}$$

Notice that the entire right diagonal of Pascal's triangle corresponds to the coefficient of y^n in these binomial expansions, while the next diagonal corresponds to the coefficient of xy^{n-1} and so on. To see how the binomial theorem relates to the simple construction of Pascal's triangle, consider the problem of calculating the coefficients of the expansion of $(x+1)^{n+1}$ in terms of the corresponding coefficients of $(x+1)^n$ (setting y = 1 for simplicity).

The rows of Pascal's triangle are conventionally enumerated starting with row n = 0 at the top. The entries in each row are numbered from the left beginning with k = 0 and are usually staggered relative to the numbers in the adjacent rows. A simple construction of the triangle proceeds in the following manner. On row 0, write only the number 1. Then, to construct the elements of following rows, add the number above and to the left with the number above and to the right to find the new value. If either the number to the right or left is not present, substitute a zero in its place. For example, the first number in the first row is 0 + 1 = 1, whereas the numbers 1 and 3 in the third row are added to produce the number 4 in the fourth row.

Binomial Expansion and Factorial Notation

The binomial theorem describes the algebraic expansion of powers of a binomial.

KEY POINTS

- According to the theorem, it is possible to expand the power $(x+y)^n$ into a sum involving terms of the form ax^by^c, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a of each term is a specific positive integer depending on n and b.
- The factorial of a non-negative integer n, denoted by n!, is the product of all positive integers less than or equal to n.
- Binomial coefficients can be written as ${}_nC_k$ and are defined in terms of the factorial function n!.

The binomial theorem describes the algebraic expansion of powers of a binomial. According to the theorem, it is possible to expand the power $(x+y)^n$ into a sum involving terms of the form ax^by^c, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a of each term is a specific positive integer depending on n and b. When an exponent is zero, the corresponding power is usually omitted from the term.

In mathematics, the **factorial** of a non-negative integer n, denoted by n!, is the product of all positive integers less than or equal to n (*Figure 8.22*). For example, $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$

The value of 0! is 1, according to the convention for an empty product. The factorial operation is encountered in many different areas of mathematics, notably in combinatorics, algebra, and mathematical analysis. Its most basic occurrence is the fact that there are n! ways to arrange n distinct objects into a sequence (i.e., permutations of the set of objects). The definition of the factorial function can also be extended to non-integer arguments, while retaining its most important properties; this involves more advanced mathematics, notably techniques from mathematical analysis.

The coefficients that appear in the binomial expansion are called binomial coefficients. These are usually written ${}_nC_k$, and pronounced "n choose k".

The coefficient of $x^{n-k}y^k$ is given by the formula:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Note that, although this formula involves a fraction, the binomial coefficient $\binom{n}{k}$ is actually an integer.

n	*n*!
0	1
1	1
2	2
3	6
4	24
5	120
6	720
7	5 040
8	40 320
9	362 880
10	3 628 800
11	39 916 800
12	479 001 600
13	6 227 020 800
14	87 178 291 200
15	1 307 674 368 000
16	20 922 789 888 000
17	355 687 428 096 000
18	6 402 373 705 728 000
19	121 645 100 408 832 000
20	2 432 902 008 176 640 000
25	$1.551\,121\,0043 \times 10^{25}$
42	$1.405\,006\,1178 \times 10^{51}$
50	$3.041\,409\,3202 \times 10^{64}$
70	$1.197\,857\,1670 \times 10^{100}$
100	$9.332\,621\,5444 \times 10^{157}$
450	$1.733\,368\,7331 \times 10^{1,000}$
1 000	$4.023\,872\,6008 \times 10^{2,567}$

Figure 8.22
Factorials
The first few and selected larger members of the sequence of factorials (sequence A000142 in OEIS). The values specified in scientific notation are rounded to the displayed precision.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/the-binomial-theorem/binomial-expansion-and-factorial-notation/

Finding a Specific Term

The rth term of any expansion can be found with the equation: $\binom{n}{r-1}a^{n-(r-1)}b^{r-1}$.

KEY POINTS

- The binomial theorem describes the expansion of a binomial: $x^{n-k}y^k = \sum_{k=0}^{n}\binom{n}{k}x^k y^{n-k}$.
- Looking at some smaller expansions, for the exponents, one can notice that the number of terms is one more than n (the exponent) and the sum of the exponents in each term adds up to n.
- Applying $\binom{n}{r-1}a^{n-(r-1)}b^{r-1}$ and ${}_nC_k = \frac{n!}{(n-k)!k!}$ one can find a particular term of an expansion without going through every single term.

The binomial theorem describes the algebraic expansion of powers of a binomial. According to the theorem, it is possible to expand the power $(x+y)^n$ into a sum involving terms of the form ax^by^c, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a of each term is a specific positive **integer** depending on n and b. When an exponent is zero, the corresponding power is usually omitted from the term.

The equation is:

$$x^{n-k}y^{k} = \sum_{k=0}^{n} \binom{n}{k} x^{k} y^{n-k}$$

Another way of writing this is ${}_nC_k = \dfrac{n!}{(n-k)!k!}$

As an example, if we go through a few expansions,

$$(a+b)^1 = a + b$$

$$(a+b)^2 = a^2 + 2ab + b^2$$

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$$

A few things can be noticed:

- The number of terms is one more than n (the exponent).
- The power of a starts with n and decreases by 1 each term.
- The power of b starts with 0 and increases by 1 each term.
- The sum of the exponents in each term adds up to n.
- The coefficients of the first and last terms are both 1 and they follow Pascal's triangle *Figure 8.23*.

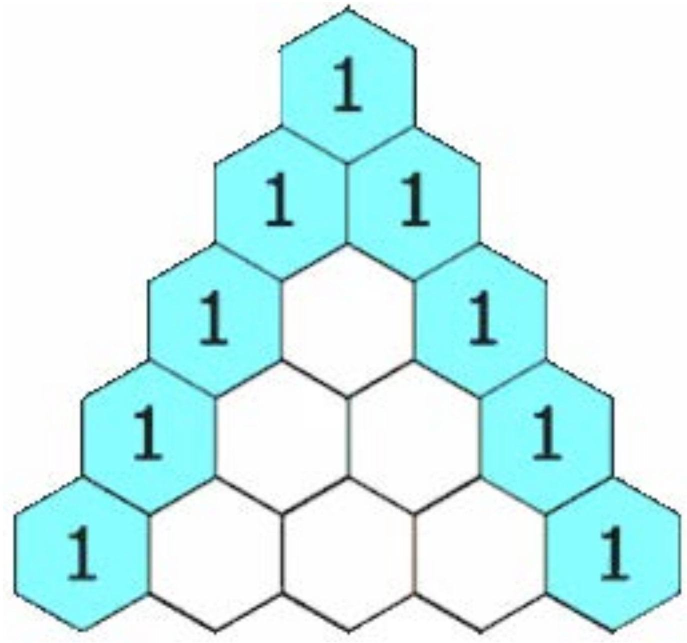


Figure 8.23
Pascal's Triangle
Each number in the triangle is the sum of the two directly above it.

If there is a short expansion, such as,

$$(x+2)^3 = x^3 + 2x^2 2^1 + 2x^1 2^2 + 2^3$$

$$(x+2)^3 = x^3 + 4x^2 + 8x + 8$$

then is easy to find a particular term. However, what about longer expansions?

Using the binomial theorem a shortcut can be made. If looking for the rth term, then:

$$\binom{n}{r-1} a^{n-(r-1)} b^{r-1}$$

Now, find the 5th term of $(3x-4)^{12}$:

$$\binom{12}{5-1}(3x)^{12-(5-1)}(-4)^{5-1}$$

$$\binom{12}{4}(3x)^{8}(-4)^{4}$$

Then, if the coefficient equation from above is applied to $\binom{12}{4}$, and the power is applied to the terms, the result is:

$$= 495 \cdot 6561x^8 \cdot 256$$

$$= 831409920x^8$$

Total Number of Subsets

The total number of subsets is the number of sets with 0 elements, 1 element, 2 elements, etc.

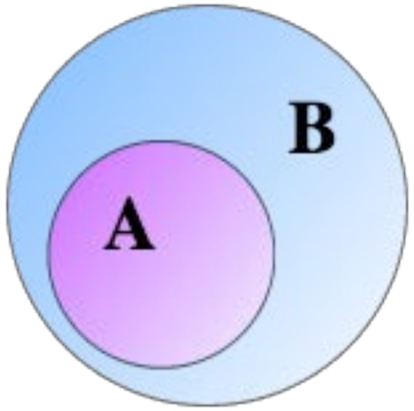
KEY POINTS

- These numbers also arise in combinatorics, where n^b gives the number of different combinations of b elements that can be chosen from an n-element set. The number of subsets containing k elements is represented by (n^k).
- According the binomial theorem, it is possible to expand the power $(x+y)^n$ into a sum involving terms of the form $a x^b y^c$, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a of each term is a specific positive integer depending on n and b.
- The total number of subsets of a set with n elements is 2^n and can be represented algebraically: $\binom{n}{0}+\binom{n}{1}+\binom{n}{2}+\ldots+\binom{n}{n}$.

The binomial coefficients appear as the entries of Pascal's triangle where each entry is the sum of the two above it. The binomial theorem describes the algebraic expansion of powers of a binomial. According to the theorem, it is possible to expand the power $(x+y)^n$ into a sum involving terms of the form $a x^b y^c$, where the exponents b and c are nonnegative integers with b + c = n, and the coefficient a

of each term is a specific positive integer depending on n and b. When an exponent is zero, the corresponding power is usually omitted from the term.

Figure 8.24 Subsets



A is a subset of B

The coefficient a in the term of $x^b y^c$ is known as the binomial coefficient (n^b) or (n^c) (the two have the same value). These coefficients for varying n and b can be arranged to form Pascal's triangle. These numbers also arise in combinatorics, where nb gives the number of different combinations of b elements that can be chosen from an n-element set. The number of **subsets** containing k elements is represented by (nk). The total number of subsets is the number of sets with 0 elements, 1 element, 2 elements, etc. This can be represented algebraically by:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n}$$

The total number of subsets of a set *Figure 8.24* with n elements is 2^n.

For example, how many subsets are in the set: {P, Q, R, S, T, U}? It has 6 elements, therefore, $2^n = 2^6 = 64 \quad subsets.$

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/the-binomial-theorem/total-number-of-subsets/

Section 7

Probability

Experimental Probabilities

Theoretical Probability

https://www.boundless.com/algebra/sequences-series-and-combinatorics/probability/

Experimental Probabilities

The experimental probability is the ratio of the number of outcomes in which an event occurs to the total number of trials in an experiment.

KEY POINTS

- In a general sense, empirical probability estimates probabilities from experience and observation.
- In simple cases, where the result of a trial only determines whether or not the specified event has occurred, modeling using a binomial distribution might be appropriate and then the empirical estimate is the maximum likelihood estimate.
- If a trial yields more information, the empirical probability can be improved on by adopting further assumptions in the form of a statistical model. If such a model is fitted, it can be used to derive an estimate of the probability of the specified event.

The **experimental probability** (relative frequency or empirical probability) is a probability that pertains to data taken from a number of trials. It is a calculable probability based not on theory, but experience. If a sample of x trials is observed that results in an event, e, occurring n times, the probability of event e is calculated by the ratio of n to x.

In statistical terms, the empirical probability is an estimate of a probability. In simple cases, where the result of a trial only determines whether or not the specified event has occurred, modeling using a **binomial distribution** might be appropriate. In such cases, the empirical estimate is the most likely estimate.

If a trial yields more information, the empirical probability can be improved on by adopting further assumptions in the form of a statistical model: if such a model is fitted, it can be used to estimate the probability of the specified event. For example, one can easily assign a probability to each possible value in many **discrete** cases: when throwing a die, each of the six values 1 to 6 has the probability 1/6. This distribution model is then created (*Figure 8.25*).

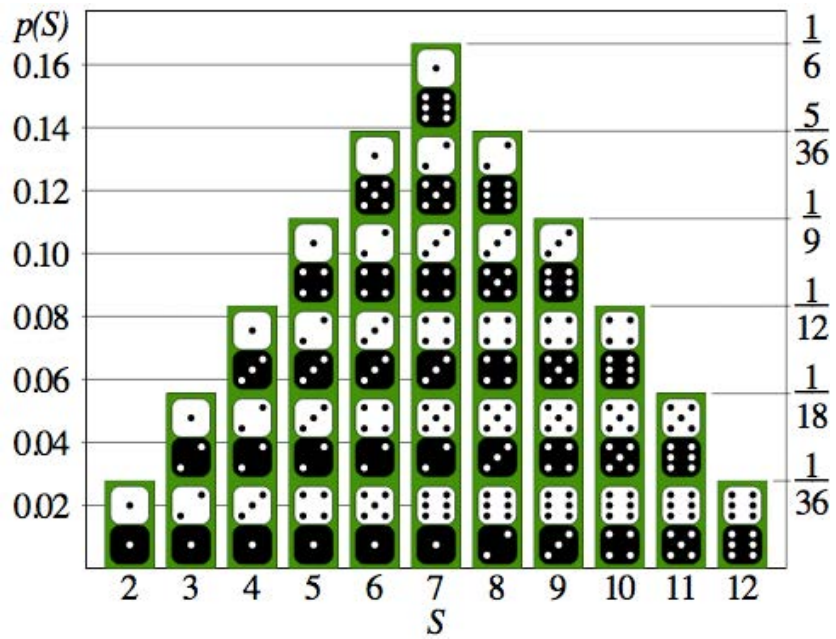


Figure 8.25 Dice Probability

Discrete probability distribution for the sum of two dice.

Advantages

An advantage of estimating probabilities using empirical probabilities is that this procedure includes few assumptions. For example, consider estimating the probability among a population of men that they satisfy two conditions:

1. that they are over six feet in height; and they are less than six feet in height.
2. that they prefer strawberry jam to raspberry jam.

A direct estimate could be found by counting the number of men who satisfy both conditions to give the empirical probability of the combined condition.

An alternative estimate could be found by multiplying the proportion of men who are over six feet in height with the proportion of men who prefer strawberry jam to raspberry jam, but this estimate relies on the assumption that the two conditions are statistically independent.

Disadvantages

A disadvantage in using empirical probabilities is that without theory to "make sense" of them, it's easy to draw incorrect conclusions. Rolling a six-sided die one hundred times, it's entirely possible that well over 1/6 of the rolls will land on 4. Intuitively we know that the probability of landing on any number should be equal to the probability of landing on the next. Experiments, especially those with lower sampling sizes, can suggest otherwise.

This shortcoming becomes particularly problematic when estimating probabilities which are either very close to zero, or very close to one. For example, the probability of drawing a number from 1 to 1000 is 1/1000. If 1000 draws are taken and the first number drawn is 5, there are 999 draws left to draw a 5 again and thus have experimental data that shows double the expected likelihood of drawing a 5.

In these cases, very large sample sizes would be needed in order to estimate such probabilities to a good standard of relative accuracy. Here statistical models can help, depending on the context.

For example, consider estimating the probability that the lowest of the daily-maximum temperatures at a site in February in any one year is less than zero degrees Celsius. A record of such temperatures in past years could be used to estimate this probability. A model-based alternative would be to select of family of probability distributions and fit it to the data set containing past years′ values. The fitted distribution would provide an alternative estimate of the desired probability. This alternative method can provide an estimate of the probability even if all values in the record are greater than zero.

Source: https://www.boundless.com/algebra/sequences-series-and-combinatorics/probability/experimental-probabilities/

Theoretical Probability

Probability theory uses logic and mathematical reasoning, rather than experimental data, to determine probable outcomes.

KEY POINTS

- Outcomes of an experiment are often equiprobable (as with heads and tails in a coin toss). In such cases, the probability of an event can be calculated logically: it is equal to the number of outcomes comprising this event, divided by the total number of outcomes in the sample space.
- If a set of choices or trials, T1, T2, T3, ..., Tk, could result, respectively, in n1, n2, n3, ...,nk possible outcomes, the entire set of k choices or trials has nk! possible outcomes. This is the Fundamental Rule of Counting.
- In the case of simple probabilities (like dice and coin tosses), outcomes are easy to count. In instances in which counting is difficult, permutations and combinations can be used to calculate outcomes.

Mathematically, **probability** theory formulates incomplete knowledge pertaining to the likelihood of an event. For example, a meteorologist might say there is a 60% chance that it will rain tomorrow. This means that in six out of every 10 times when the world is in its current state, it will rain. This probability is

determined through measurements and logic, but not through any experimental findings (the future has not yet happened). As such, the meteorologist's 60% verdict is a theoretical probability, and not the result of any proven experiment.

Often, in experiments with finite sample spaces, the outcomes are **equiprobable** (as with heads and tails in a coin toss). In such cases, the probability of an event can be calculated logically: it is equal to the number of outcomes comprising this event, divided by the total number of outcomes in the sample space.

For example, the probability of rolling any specific number on a six-sided die is one out of six: there are six, equally probable sides to land on, and each side is distinct from the others. If the six on the die were changed to a one, you could logically conclude that the probability of rolling a one would be two out of six (or one out of three). This is a theoretical probability; testing by rolling the die many times and recording the results would result in an experimental probability.

While counting outcomes may appear straightforward, it is in many circumstances a daunting task. For example, consider the number of distinct subsets of the integers {1, 2,..., n} that do not contain two consecutive integers. This number is equal to:

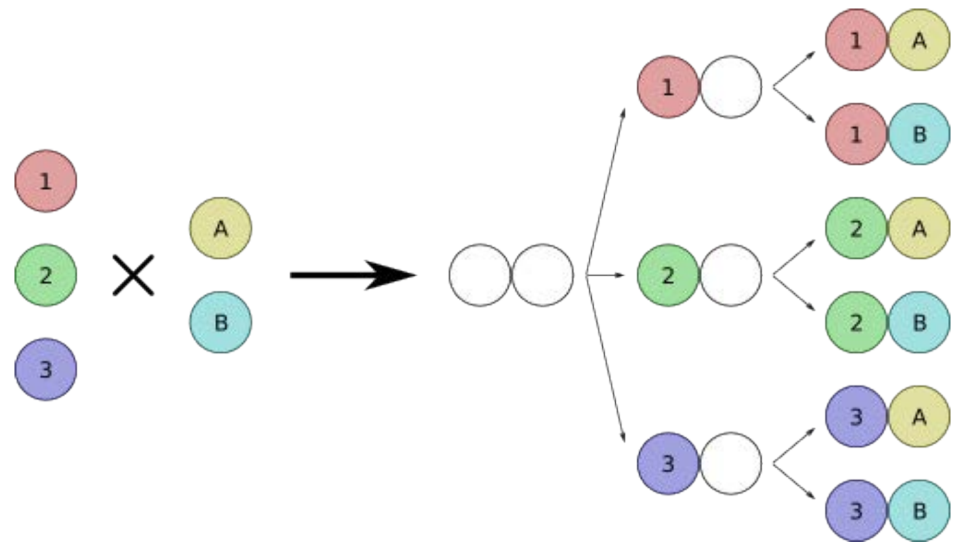
$$\frac{\phi^{n+2} - (1-\phi)^{n+2}}{\sqrt{5}}$$

where $\phi = (1+\sqrt{5})/2$ is the **golden ratio**. It can also be obtained recursively through the Fibonacci recurrence relation. Calculating the number of ways that certain patterns can be formed is part of the field of combinatorics.

The Fundamental Rule of Counting

If a set of choices or trials, T_1, T_2, T_3, ..., T_k, could result, respectively, in n_1, n_2, n_3, ...,n_k possible outcomes, the entire set of k choices or trials has n_k! possible outcomes. By the Fundamental Rule of Counting, the total number of possible sequences of choices

Figure 8.26 The Counting Principle



By the Fundamental Rule of Counting, the total number of possible sequences of choices is a permutation of each of the items.

is 5×4×3×2×1 = 120 sequences. Each sequence is called a permutation (or ordering) of the five items (*Figure 8.26*).

Permutations

A permutation is an arrangement of unique objects in which order is important. In other words, permutations using all the objects: n objects, arranged into group size of n without repetition, and order being important. The number of possible permutations of a set size of n in which k elements are drawn can be calculated by:

$$\frac{n!}{(n-k)!}$$

Combinations

A combination is an arrangement of unique objects, in which order is not important. For example, the number of possible combinations of n objects, arranged in groups of size r can be calculated by:

$$\frac{n!}{r!(n-r)!}$$